TRACKER STUDIES

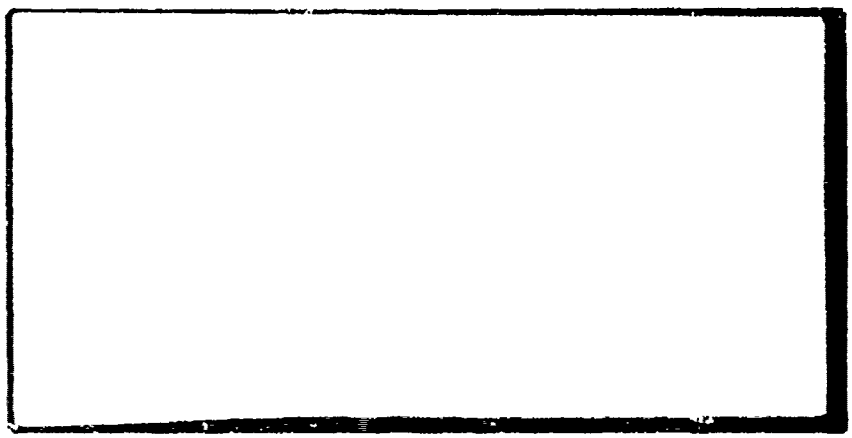
R. N. DeWitt

ITT Electro-Physics Laboratories, Incorporated

Prepared for:

Office of Naval Research

June 1975



PROJECT REPORTS

Normally contractually required reports describing completed phases of customer-supported project; may be used to mark completion of any major investigation, however. **(Green Cover)**

ELECTRO-PHYSICS LABORATORIES INC.

International Telephone and Telegraph Corporation • Defense - Space Group

9140 Old Annapolis Road, Columbia, Maryland 21045

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CHAPEL BELL No. 115	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TRACKER STUDIES		5. TYPE OF REPORT & PERIOD COVERED Final Report, Feb.-Dec. 1974 6. PERFORMING ORG. REPORT NUMBER Project Report No. 278
7. AUTHOR(s) Dr. R.N. DeWitt		8. CONTRACT OR GRANT NUMBER(s) N00014-74-C-0106
9. PERFORMING ORGANIZATION NAME AND ADDRESS ITT Electro-Physics Laboratories, Inc. 9140 Old Annapolis Road Columbia, Maryland 21045		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62751 N, Task Area RF12-141-001
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Field Projects Programs, Code 464 Arlington, Virginia 22217		12. REPORT DATE June 1975 13. NUMBER OF PAGES 140
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Same		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A
16. DISTRIBUTION STATEMENT (of this Report) Unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) Unlimited		
18. SUPPLEMENTARY NOTES N/A		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Multi-Mode HF Propagation Predictive Filters Kalman Filter Mode Identification HF Radar Tracker		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The existence of multi-mode propagation can have two deleterious effects on an OTH radar tracker; it can lead to ambiguity in the interpretation of the range of detected targets, and it can lead to difficulties in determining their number. Because the identity of a mode cannot be uniquely determined from a single detection sample, any testing to determine mode identity must be performed at or beyond the tracker.

Three methods are considered for identifying propagation modes; one based on testing for the occurrences of echoes in regions or combinations of regions of

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (continued)

the radar-observable space accessible to only a single propagation mode, another based on the amplitude fluctuations observed on echoes, and a third in which mode separations are tracked and taken into account.

Predictive filters are given general treatment. The Kalman filter tracking algorithm is discussed with the modifications required to implement the first and third (above) mode-identification techniques. The use of signal strength fluctuations as a mode-identification criterion is examined and dismissed as being insufficiently predictable to provide a reliable means for identifying propagation modes.

ITT-EPL PROJECT REPORT NO. 278

CHAPEL BELL REPORT NO. 115

TRACKER STUDIES

This report is the ONE-HUNDRED-FIFTEENTH of a series of CHAPEL BELL reports (formerly TEPEE reports) made under Contract N00014-74-C-0106 dated 1 July 1973 under the sponsorship of the Office of Naval Research, Code 464.

ITT-EPL PROJECT REPORT NO. 278

CHAPEL BELL REPORT NO. 115

TRACKER STUDIES

Date: June 1975

Prepared by R N DeWitt

R. N. DeWitt
Manager, Advanced Systems

Approved by

R. Wagner
Program Manager

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST (Continued)

ABSTRACT

The existence of multi-mode propagation can have two deleterious effects on an OTH radar tracker; it can lead to ambiguity in the interpretation of the range of detected targets, and it can lead to difficulties in determining their number. Because the identity of a mode cannot be uniquely determined from a single detection sample, any testing to determine mode identity must be performed at or beyond the tracker.

Three methods are considered for identifying propagation modes; one based on testing for the occurrences of echoes in regions or combinations of regions of the radar-observable space accessible to only a single propagation mode, another based on the amplitude fluctuations observed on echoes, and a third in which mode separations are tracked and taken into account.

Predictive filters are given general treatment. The Kalman filter tracking algorithm is discussed with the modifications required to implement the first and third (above) mode-identification techniques. The use of signal strength fluctuations as a mode-identification criterion is examined and dismissed as being insufficiently predictable to provide a reliable means for identifying propagation modes.

1. INTRODUCTION

(U) The investigations reported herein were conducted during the period February through December 1974 under the sponsorship of the Office of Naval Research, Code 464, under Contract N00014-74-C-0106.

1.1 The Role of Mode Identification in an OTHR System

To permit OTH radar data to be interpreted and displayed in geographic coordinates, a coordinate conversion process is required. Because the conversion process must be done differently from one mode to another, it is necessary to correctly identify the propagation mode by which each target is detected. In some cases, this is no problem because by properly choosing frequency and other operating parameters of the radar, it is possible to assure that targets are illuminated via a unique mode. However, all too often such a solution is precluded, either by the impossibility of managing propagation so as to isolate a single mode, or simply because conflicting considerations make such a choice of operating parameters inadvisable.

In particular, propagation management is already greatly constrained by the need to provide adequate illumination in the surveillance area, and it would be undesirable to compromise the amount of energy placed on target for other considerations. The methods to be discussed later are intended to avoid, or at least to minimize, the need for such a compromise.

There is no practicable way in which the mode of propagation can be determined from a single pulse or single doppler dwell observation of a target. Although in principle such a determination could be made by measuring the vertical angle of arrival of the echoed radiation, in practice this would require greater antenna resolutions than are economically feasible. Most of the methods to be considered here involve multiple observation of the target--that is, observations made during different pulses or doppler dwells of the radar. Because such multiple observations of a target are not associated until the "tracker" stage of the radar data processing, the methods to be discussed here are, of necessity, applied at or beyond the tracker. One technique that has been considered can be applied within a single frame of data but can be applied to tracked data as well. Therefore, the discussion

will hereafter assume that tracked data are to be used. In such a mode of operation, the tracker operates on data in terms of the radar observables--time delay of the echo, doppler shift of the echo, and angle of the echoed radiation relative to a reference. The only requirements imposed on the tracker are that it be tolerant to any discontinuities of the track and of apparent accelerations of the targets produced by the propagation characteristics of the ionosphere.

1.2 The Problems of Ambiguous and Multiple Returns

For the purpose of the present discussion, it is useful to identify two parts of the problem caused by the radarist's uncertainty regarding the mode of propagation. In order to correctly interpret the data from an OTH radar, it is necessary to perform a one-to-one mapping of the target's position in radar observable space to geographical coordinates. However, under multiple mode conditions, this mapping process is complicated in two ways. First, the detected target may occur at a point in the radar observable space that can be mapped to any of two or more points in the true geographic space, depending upon what mode is assumed to have supported the propagation. Until the responsible mode can be identified, it is necessary for the radarist to carry along a set of multiple hypotheses regarding the position of the detected target. In effect, the data must undergo a one-to-many mapping from the radar-observables space to geographic coordinates. We here speak of this as the mode ambiguity problem, the effects of which can be stated:

1. The true number of resolved targets is readily apparent.
2. But the true position of the target is one of perhaps several discretely determined positions.

The second kind of problem occurs if the target happens to be at a point in the space of geographical coordinates at which it is illuminated via more than one propagation mode. The effect of this will be that the target will manifest itself at several points in the radar-observable space. A distinct return from the target can occur for each propagation mode. Additional returns will also arise corresponding to outward propagation via one mode and return propagation via another. However, mixed-mode returns will occur with the same delay regardless of the order in which any given pair

of modes occur--whether, for example, an E mode occurs in the outward leg and an F mode on the return or vice versa. In view of the above, we see that the number of returns from a target in a region of N modes of illumination will be equal to the number of ways N elements can be drawn twice, with replacement, to form unordered pairs. Thus, the number of returns, n, is:

$$n = \sum_{i=1}^{N} i$$

Unless the radar has resolution on the order of 1 millisecond, the ordinary and extraordinary modes of propagation will remain unresolved and manifest themselves only as a polarization rotation and associated fading. Propagation via the F layer will occur in two modes--the high ray and the low ray. In addition, there exists the possibility of propagation via the E layer and F1 layers when they exist. Depending on the height and electron densities of the layers, various combinations of these may propagate radar illumination to the region of surveillance and give rise to the multimoding effects. Note that the effect of multimoding differs from that of mode ambiguity in that:

1. The true number of targets cannot be determined until an association is made among the multiple echoes produced by each individual target.

2. Knowledge of the true position of the target is not precluded by multiple moding. Only if one or more of the returns lies in ambiguous regions of the radar-observable space will there be range ambiguity of the kind discussed earlier.

The practical effect of multimoding is that it may produce erroneous estimates of the number of targets detected in the surveillance region. A correct coordinate conversion from radar observable space to geographic coordinates would lead to coincident points, readily interpreted as being a single target. However, all too often, errors in the coordinate conversion process will lead to imperfect coincidence and difficulty in interpretation.

A second practical implication of the multiple mode effect is that the multiple returns could be combined non-coherently, or perhaps even coherently, to improve the detection sensitivity of the radar. However, to

do this, it would be necessary to convert from radar observable coordinates to geographical coordinates prior to tracking, and various hypothetical tracks would have to be carried for radar returns occurring in ambiguous regions of the radar observable space. Consequently, although better performance could be expected with such an arrangement, it would be at the expense of more powerful data processing. The coordinate conversion processor would have to be powerful enough to perform a conversion for each ambiguity of each detection, and the tracker would have to be powerful enough to carry the additional hypothetical tracks that arise from the mode ambiguity.

1.3 Mode Identification and ECM

Many of the considerations involved with mode identification also have applicability to the problem of making the radar immune to deceptive forms of electronic countermeasures. In order to produce false targets in an OTH radar, it is necessary to transmit a false pulse delayed with respect to the time of transmission of the radar illumination pulse by an amount appropriate to the range at which the false target is to appear. However, a sophisticated radarist employing mode identification techniques would not depend only upon the delay of the radar signal as an indication of the apparent target range. He would also seek other cues, such as the maximum frequency illuminating the target via a particular mode, and the degree of coincidence among positions of a target as deduced from multiple modes of coverage. The need for sufficient information to simulate all of these effects would make the task of the radarist's ECM opponent much more difficult, and were proper measures taken on the part of the radarist, vulnerability to spoofing ECM could nearly be precluded.

Consider a radar operating with a nonrepetitive, unpredictable waveform--such a radar is susceptible only to false targets at delays greater than that of the spoofing transmitter. (For the moment, we assume that we are dealing only with a single propagation mode.) However, if the radarist increases frequency until the point is reached at which the spoofing transmitter (and receiver) is in the skip zone, while the apparent position of the target is still being illuminated, the spoofing can no longer be carried on because the spoofer becomes decoupled from the radar. The ability to keep track of

tests of this kind is a large part of the mode identification process, so the ability to recognize spoofing ECM is an additional benefit to be gained by employing the techniques discussed here.

1.4 Fixed Frequency Versus Multiple Frequency Operation

Many of the techniques here depend on the radar having some degree of frequency agility. For our purpose, it will usually be satisfactory for the radar to be able to change frequency from one doppler dwell to the next. The changing of frequency can either be carried out in a predetermined pattern (e.g., by assigning frequencies by means of a pseudo-random sequence) or can be made adaptive to the set of echoes of the moment. The latter procedure would be the preferred one if the frequencies required to identify propagation mode were unsatisfactory for general illumination of the surveillance area. However, such a procedure would require that the tracker-mode identifier be given the capability of selecting the radar's operating frequency.

If the radar is constrained to fixed-frequency operation, the ability of the radar to carry out tests to determine propagation modes is considerably limited. Tests involving determination of the maximum observable frequency to a target then reduce to observations of the radar delay at which the target enters the mode's skip zone. Because this kind of test is not likely to be generally satisfactory, consideration was given to a second set of mode identification tests based on the time variation of the amplitude of the radar echo from the target. A discussion of this possibility is to be found in Section 4.

2. THE PROBLEMS OF MODE AMBIGUITY AND MULTIPLICITY

2.1 Prevalence of Multimode Conditions

As mentioned in Section 1, it is often possible to take steps in the management of the OTH radar to reduce the number of modes with which the radar must contend. A good example of this is the practice of restricting the radar's operation to radiation propagated via the F modes. This is done by choosing a sufficiently high operating frequency so that the waves penetrate the E layer except at very low radiation angles. At these low angles the combined effects of reduced antenna gain and D-layer absorption largely reduce the radar's response to clutter and targets at ranges of interest.

Unfortunately, there are many circumstances in which this is not possible; for example, when the critical frequencies of two layers are not sufficiently different. Even when the effects of other ionospheric layers can be avoided, it is still necessary to contend with the various forms of multimoding within the layer. The occurrence and characteristics of the various forms of multimoding are discussed next for background.

2.1.1 High Ray-Low Ray Propagation

Each ionospheric layer supports radio propagation between a pair of points via several modes. Some of the multiplicity of modes are the result of the earth's magnetic field on propagation. However, ignoring for the moment the mode splitting attributable to magneto-ionic action, we are left with two propagation modes that are supported by each layer: the low angle ray and the high angle or Pedersen ray. It is the low ray that is normally counted on to provide the propagation from radar to target, being more stable and not subject to the defocusing action to which the Pedersen ray is subjected.

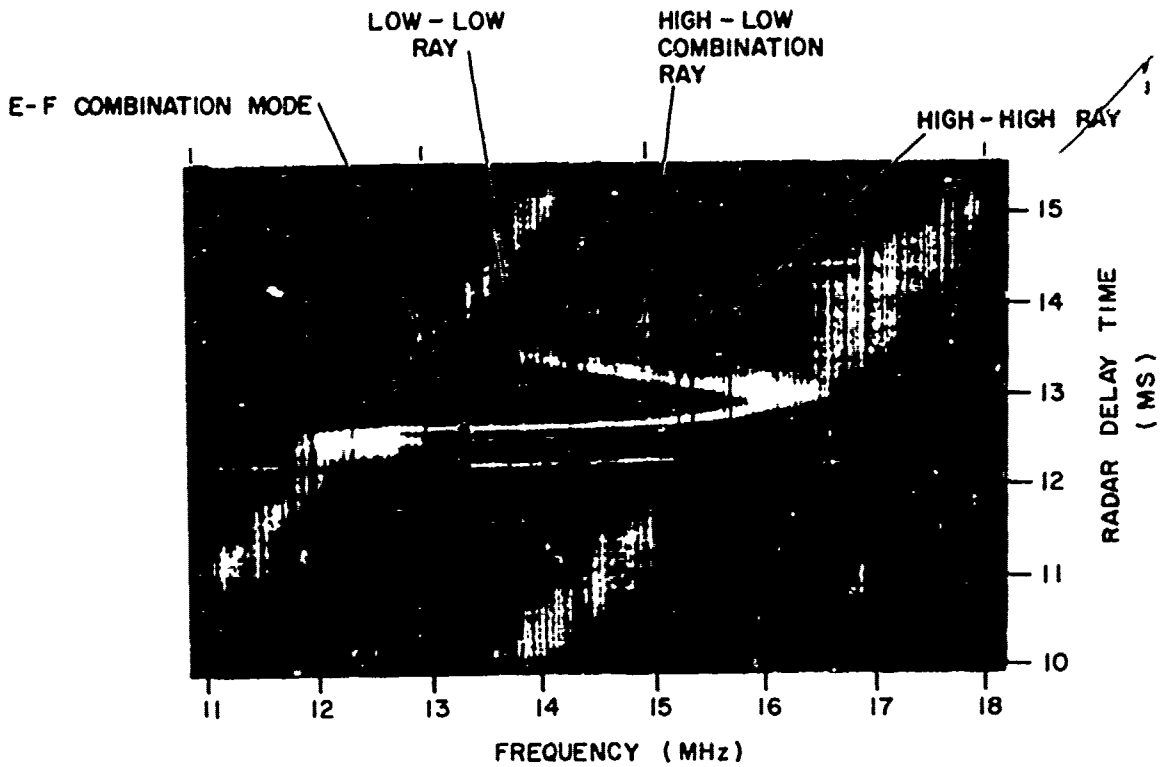
The high and low ray modes of propagation can produce three signals from each target: one trace produced by echoes propagated by the low ray mode, a second produced by high ray propagation, and a third, at intermediate delay, which is produced by echoes that have propagated by high ray in one direction and low ray in the other.

In this case, not only is there a multiplicity of modes but also an ambiguity because each trace lies in a region of the frequency-delay plane that is accessible to the other two signal paths. Although the association of the traces is readily apparent on an F scan, where they join at the layer's skip-zone caustic, their association would not be as readily apparent in a single radar observation or in a series of observations at a single frequency.

For strong echoes, such as those from a repeater, two or more of these modes can often be seen to extend several megahertz below the junction frequency. However, the high ray is subjected to an appreciable defocusing effect, causing signals propagated by that mode to be substantially weaker than those propagated via the low ray, particularly as the operating frequency departs more and more from the junction frequency for the target's range.

Tornatore (1972) of ITT-EPL has measured the relative strength of signals propagated one way via the high and low rays and has found that the low ray signal strength exceeded that of the high ray by median ratios of 3, 6 and 10 db as measured 1, 2, and 3 MHz below the junction frequency, respectively. These results can be applied directly to the two-way radar case to give the relative amplitude of the low and mixed-ray modes but must be doubled to give the ratio of low-to-high ray radar signal strength.

Figure 2-1 shows an oblique ionogram indicating radar propagation via the F layer between Whitehouse, Virginia, and Lincoln, Nebraska. The regions of high, low and mixed-mode ambiguity are indicated in the lower tracing.



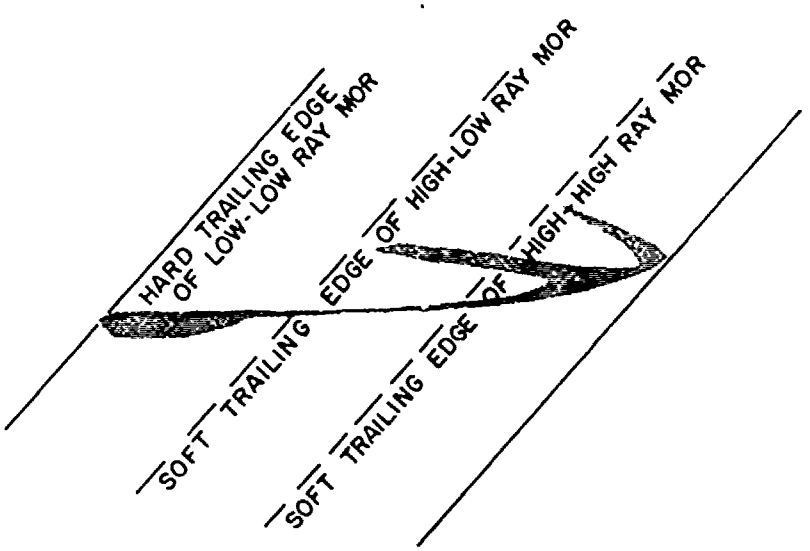
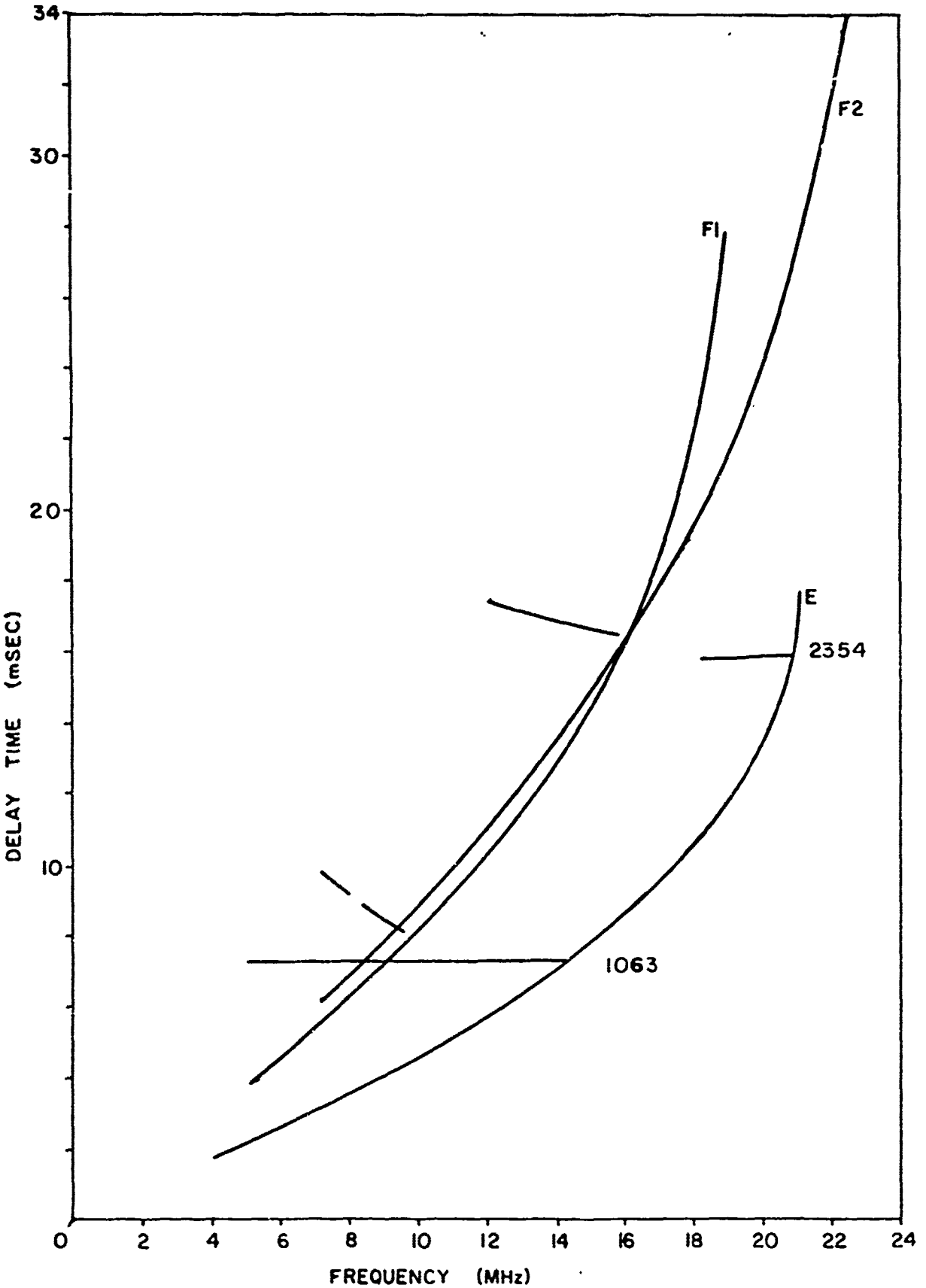


Figure 2-1. Tracing Showing Boundaries of MOR's Caused by High Ray-Low Ray Multimoding

2.1.2 Multiple Layer Multimoding

Under daytime conditions the ionosphere frequently comprises two or more layers capable of supporting propagation; and even under nighttime conditions, during spring and summer months, a sporadic E layer frequently accompanies the F layer and produces multimoding.

Depending on the distribution of electron densities in the layers, either the F mode or an underlying mode is capable of supporting propagation to a given distance at higher frequencies. The anomalously high F layer critical frequency occurring on winter days are typical of those cases where the F layer supports the higher frequency of propagation, permitting undesired modes to be avoided by selecting a suitably high operating frequency. On the other hand, during summer days, the opposite condition exists and the E layer often supports the higher frequency of propagation to any range. Figure 2-2 shows the computed traces of leading-edge-focused ground clutter superimposed on oblique ionograms to two ges. These traces were computed for a model ionosphere appropriate summer daytime propagation southward from the ONR, Whitehouse, Virginia, site. Note that the E layer is shown to support propagation to radar delays of 8 to 16 ms at frequencies some 4 MHz higher than does the F layer. However, the radiation angles required to achieve the E layer propagation are quite low, being about seven degrees for 1000 km propagation (≈8 ms) and going to zero degrees before the 16 ms delay can be achieved. The reduced gain of practical antennas at such low radiation angles usually prevents E-mode propagation except to relatively short ranges. Consequently, in these cases the E layer does not directly contribute to the mode ambiguity and multimoding problems. However, it does have a serious indirect effect-- it obscures the higher layers over a broad range of frequencies. This effect is shown in Figure 2-2 by the abruptly truncated low ray traces. The result is that the radar must be operated at a frequency near the MUF of the F layers. This leads not only to the necessity to cope with the high ray propagation but



1736

Figure 2-2. Traces of Leading-Edge-Focused Ground Clutter and Discrete Targets at Ranges of 2354 and 1063 KM, Computed from a Model Ionosphere Typical of a Summer Daytime. Pronagation southeast from Whitehouse, Virginia, is assumed.

also with a multimoding problem involving F1 and F2 layer propagation. Figure 2-3 is a skywave backscatter record made with a large cross section repeater in the field of view. Note that the repeater trace, which has been produced by F layer propagation, does not have a low-ray component extending to low frequencies. Rather, it is limited to the frequencies near the MOF where the high ray and mixed ray produce multimoding.

2.1.3 Sporadic E Modes

During the spring and summer (near summer solstice) months, an additional ionospheric component compounds the multimoding problem further. The E layer is subject to localized enhancements leading to critical frequencies that can be in excess of 7 MHz. Besides intensifying the obscuration problem posed by the normal E layer during the daytime hours, sporadic E can also introduce similar problems at night. The probability of its occurrences with critical frequencies in excess of 5 MHz is on the order of 30%. Figure 2-4 shows a skywave backscatter record in which sporadic E is strongly evident at ranges out to about 10 msec. In these cases, propagation via the sporadic E layer, although limited and uncertain, can be managed to avoid multimoding effects. The same cannot be said of F-mode propagation. Simple sporadic E modes will not pose a problem at radar delays in excess of 10 msec., but multiple hop modes are a possibility.

The likelihood of two-hop E_s modes is reduced by the requirement that sporadic E must occur at two independent reflection points. However, if the operating frequency is chosen too far below the F layer MOF for the range of interest, the possibility of two-hop modes involving the E and F layer arises. The probability of these occurring is approximately the same as for the one-hop E mode, since only one E-layer reflection is required.

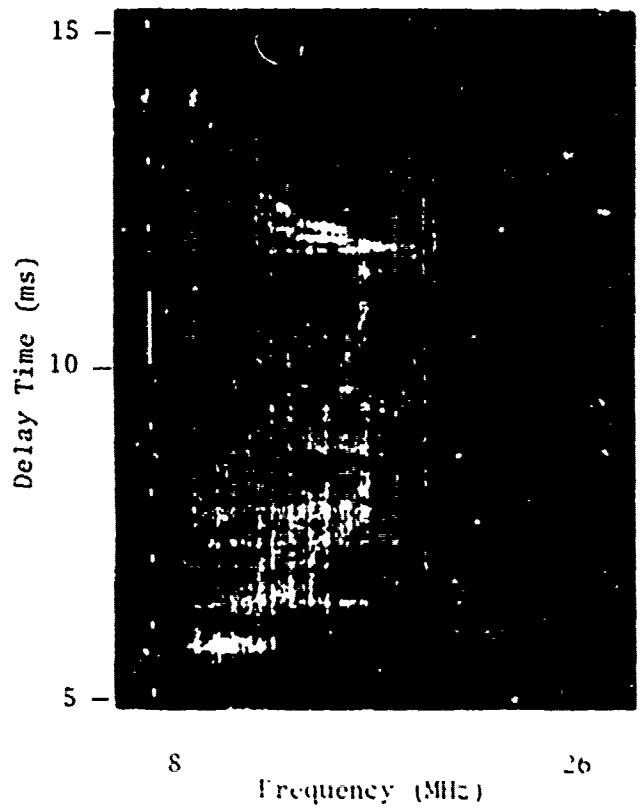


Figure 2-3. Skywave Backscatter f-Scan Showing Echo Supported by F-Layer Propagation Which is Confined to Frequencies Near the MUF Because of Underlying Obscuring Layers

11 August 1971
2258 Z
Transmit Steer - 290°
Receive Steer - 288.5°

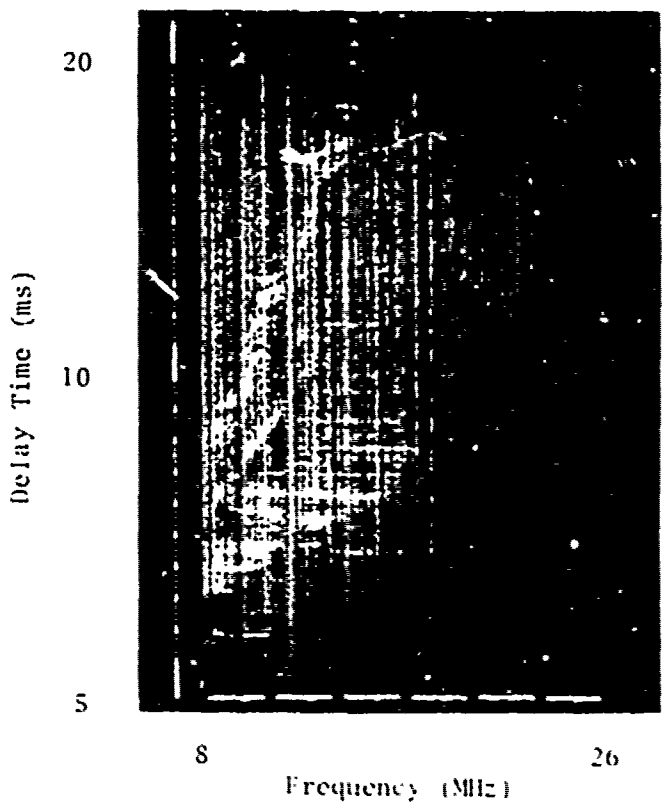


Figure 2-4. Skywave Backscatter f-Scan in Which Echoes Caused by Sporadic-E Propagation Appear Prominently at Radar Delays Approaching 10 Msec

2.1.4 Multihop Modes

Because multihop modes require ionospheric reflections at angles more nearly vertical than do single hop modes, there is usually an interval of frequency at which the single hop mode is supported but which penetrates the layer at the multihop angle. Figure 2-5 shows an example of a one-hop mode, a two-hop mode, and the intermediate mixture mode. Note that the mixture mode and the two-hop mode both reach an upper frequency limit of about 8 MHz lower in frequency than does the one-hop. This interval would provide ample scope for avoiding multimoding problems.

2.1.5 Unresolved Modes

In addition to the multiple moding effects described above, there are other forms of multimoding. The splitting of propagated waves into two characteristic magneto-ionic modes is an important one of these. Depending upon the geometry of the propagation path and the geomagnetic field, these modes can differ by as much as a MHz in the MOF to a given range. The difference in radar delay associated with magneto-ionic multimoding is of order a few microseconds for typical OTH radar propagation, so it remains unresolved unless the radar has a bandwidth that is a substantial fraction of a MHz. If unresolved, this multimoding manifests itself as a fading effect observed in echoes from linearly polarized scatterers.

Other forms of multimoding have been invoked to explain other forms of echo fading and dispersion in time. Usually,their effect is only to induce signal fading. This will be discussed further in a later section where the use of fading rate as an identifying characteristic of propagation mode is considered.

2 November 1971
1412 Z
Transmit Steer - 290°
Receive Steer - 288.5°

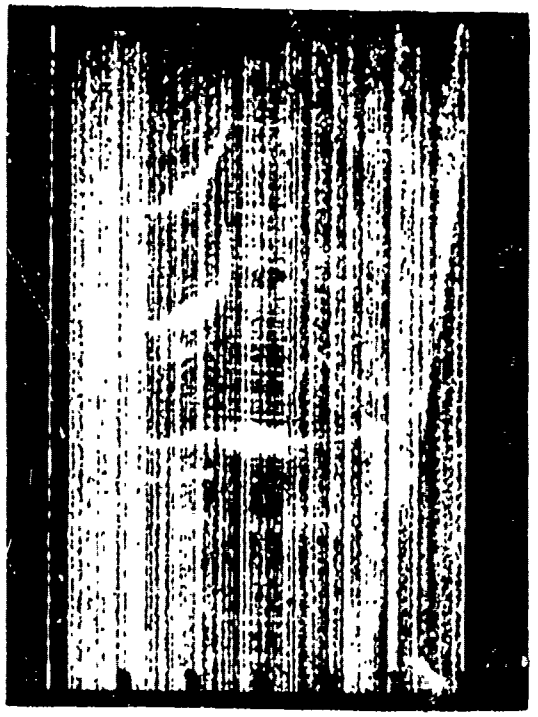


Figure 2-5. Skywave Backscatter f-Scan in Which Multiple-Hop-Propagated Echoes from a Repeater Appear Prominently. The upper trace is caused by two-hop propagation to and from the repeater. The middle trace is caused by one-hop propagation in one direction and two-hop in the other.

2.2 Mode Regions and Boundaries

It is useful in the ensuing discussion to define some terms and concepts for the special purpose of dealing with the OTH radar mode identification problem. At this point we introduce two; these are the mode's observable region and the mode's coverage region. Each propagation mode has associated with it one of each kind of these regions. The MOR (Mode-Observable Region) is defined in the radar delay-frequency plane and is the locus of all points on that plane at which a radar echo can appear as the result of energy that has propagated via that particular mode and has been scattered by a target at the earth's surface. The latter proviso can and must be made in order to limit the region to that in which targets of interest can occur. Aircraft fly sufficiently near to the surface to produce echoes only in the mode-observable region, while meteors and space vehicles can produce echoes outside the region.

The second term, mode-coverage region, is used to denote a region in the frequency-distance plane which represents the space covered by the radar at any frequency, using a particular propagation mode in its operation. The mode-coverage region can be derived from the mode-observable region by applying the coordinate conversion process, appropriate to that mode, to the radar delay coordinate. However, because the relationship of radar delay to ground range differs from one mode to another, the relationship among the various mode-observable regions (for different propagation modes) will not be the same as that among the mode-coverage regions. For example, an overlapping of two mode-observable regions does not necessarily imply a corresponding overlapping of the same modes' coverage regions. In the discussion thus far, the radar observable-angle with respect to the array and the spatial coordinate-bearing from the radar have not been mentioned. However, the structure of the ionosphere and, consequently, the required conversion process between radar delay and ground range can vary with azimuth. Were the variation rapid enough, it would be necessary to define the mode-observable and mode-coverage regions in a three-dimensional space. We here adopt the usual practice of assuming that the effect of azimuthal variation can be handled adequately by simply redefining the mode-observable and mode-coverage regions for each angle from the antenna.

2.2.1 Types of Boundaries

As noted above, the delay time-frequency plane can be partitioned into mode-observable regions, each a characteristic of a particular propagation mode. The overlapping portions are defined as ambiguous regions. The location of the regions' boundaries in the delay time-frequency domain may be established by observation or by derivation from a model ionosphere.

Segments of the boundary encompassing the regions fall into two categories, hard or soft. Hard boundaries are those which are caused by well-defined geometrical limiting effects such as:

a. The leading edge of ground backscatter.

b. The caustic focusing region.

c. The trace of the minimum propagation elevation angle in the delay time-frequency domain for the lowest ionosphere layer.

d. The focusing at maximum range (Croft, 1967).

The hard boundaries can be directly computed by ray tracing techniques. With the exception of the minimum elevation angle of the lowest layer, the remaining boundaries are often readily observed on backscatter records as enhancements of echo strength. Figure 2-6 illustrates an example of hard boundaries. An ambiguity region is delineated by the leading edge of the E-mode backscatter and the focusing at the maximum range of the F layer.

The soft boundaries are those boundaries at which propagation is geometrically possible, but at which the noise or path loss has degraded the signal-to-noise ratio to a degree such that the probability of detection is too small to be useful. The maximum delay time trace on a backscatter record corresponding to the minimum elevation angle for an F1 or F2 layer (E layer present) is an example of soft boundary. System effects such as antenna gain, equipment loss or gain and ionospheric loss mechanisms modify the trace of a soft boundary. Figure 2-7 illustrates an example of a soft boundary. Figure 2-8 shows hard and soft boundaries established by ray-tracing in a complex ionosphere composed of three superimposed quasi-parabolic layers. The leading edge of the E, F, and F2 layers are computed

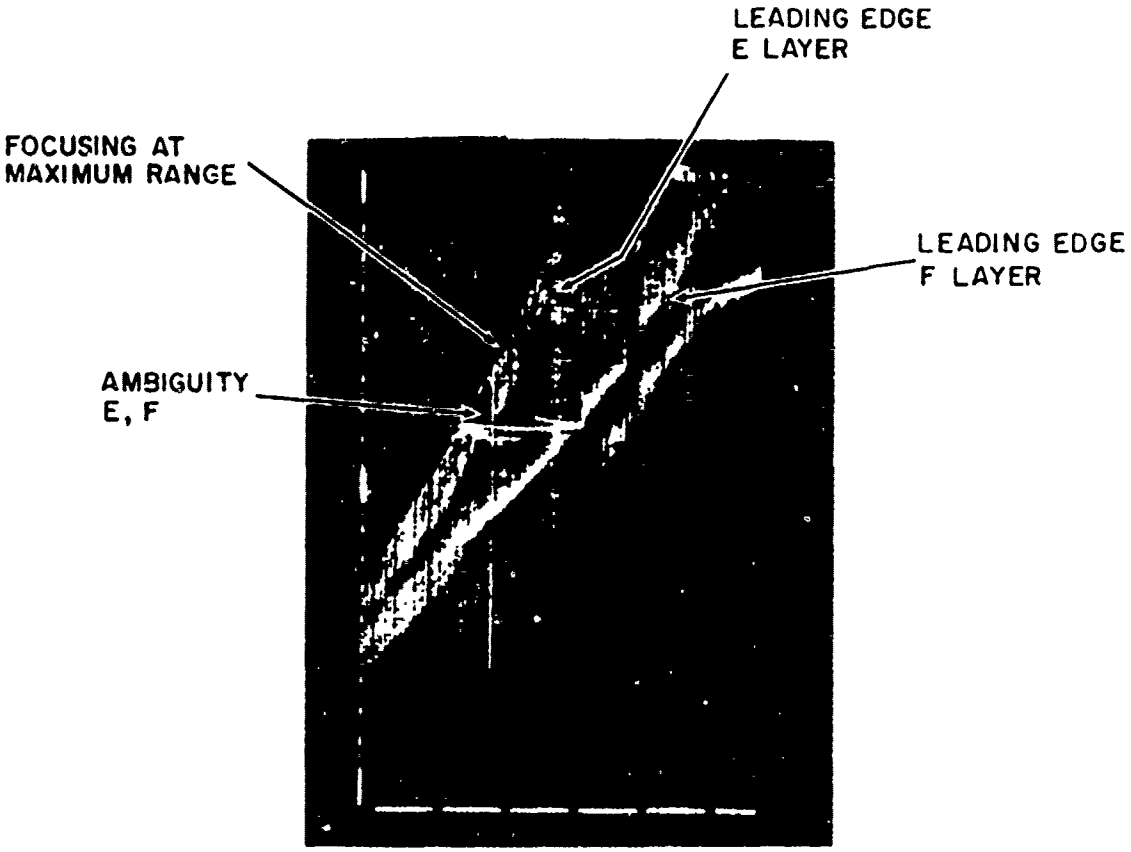


Figure 2-6. Backscatter Trace Illustrating Hard Boundaries and an Ambiguity Region

1 - One-hop repeater trace
2 - Mixed one-hop/two-hop repeater trace
3 - Two-hop repeater trace

15 May 1969
1342 Z

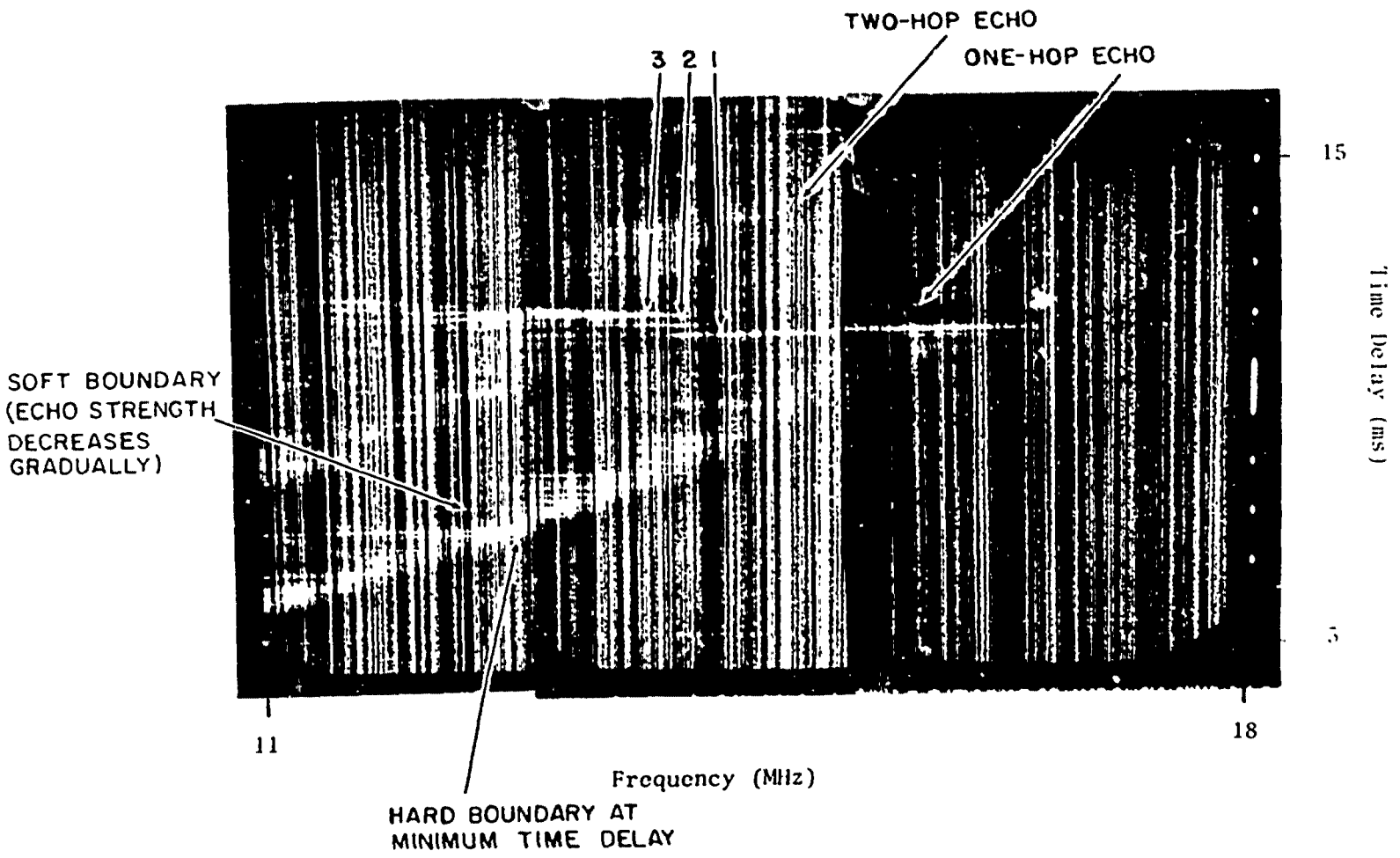


Figure 2-7. Skywave Backscatter f-Scan Showing a Soft Boundary at the Trailing Edge of a Mode Observable Region

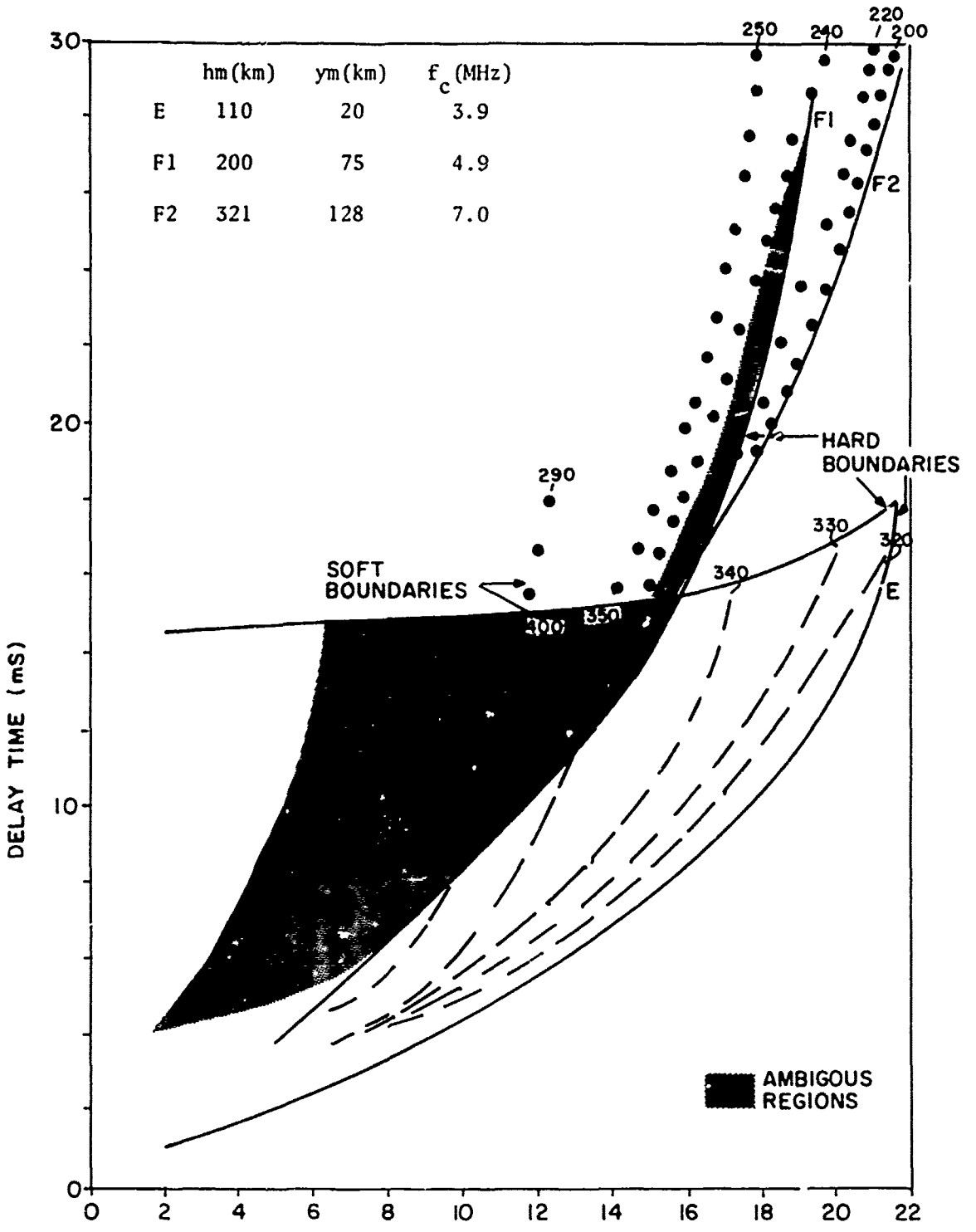


Figure 2-8. Ambiguous Regions Established by Hard and Soft Mode Boundaries

Antennas with little gain near the horizon will produce a soft boundary at, for example, the longer time delay boundary of the E layer mode-observable region. This effect was not considered in making this diagram, so this boundary is indicated to be a hard one.

5-172

by finding the corresponding delay time trace where the condition

$$\frac{d\tau}{d\alpha} = 0$$

where τ = delay time

α = elevation angle

holds. The zero degree ray establishes a segment of the E mode boundary. The trailing edge soft boundaries may be established with the aid of system sensitivity contours such as that shown in the figure. The constant loss contours are superimposed on the E and F regions. These loss contours follow the general shape of the leading edge of the clutter, and the loss decreases to a minimum value near the leading edge. This decrease in echo strength corresponds to the fadeout at lower frequencies which occurs in Figure 2-7. The soft boundary is the locus of points in the frequency-radar delay plane at which the signal-to-noise ratio is only marginally adequate for useful observation.

In the above example, the soft boundary at the low frequency (or long delay) edge of the mode-observable regions was imposed by propagation path loss that reduced signal strength to a degree at which it became unusable. We may now consider another example, in which a soft boundary is produced by another mechanism.

Figure 2-9 is a computed skywave backscatter trace for a summer daytime sounding in the southward direction from Whitehouse. The caustic-focus traces of the F1 and F2 layers are nearly coincident, while that of the E layer lies about 6 MHz higher in frequency than do the F layer traces. Traces for what would be discrete targets at range of 2354 km and 1063 km are also shown. This figure indicates the high frequency "hard" boundaries imposed by the geometrical optics of the situation, but it does not indicate other boundaries of the mode-observable regions. Figure 2-10 shows a plot in the elevation angle-radar delay (α-τ) plane for rays of a 14-MHz signal. This diagram indicates the obscuring effect

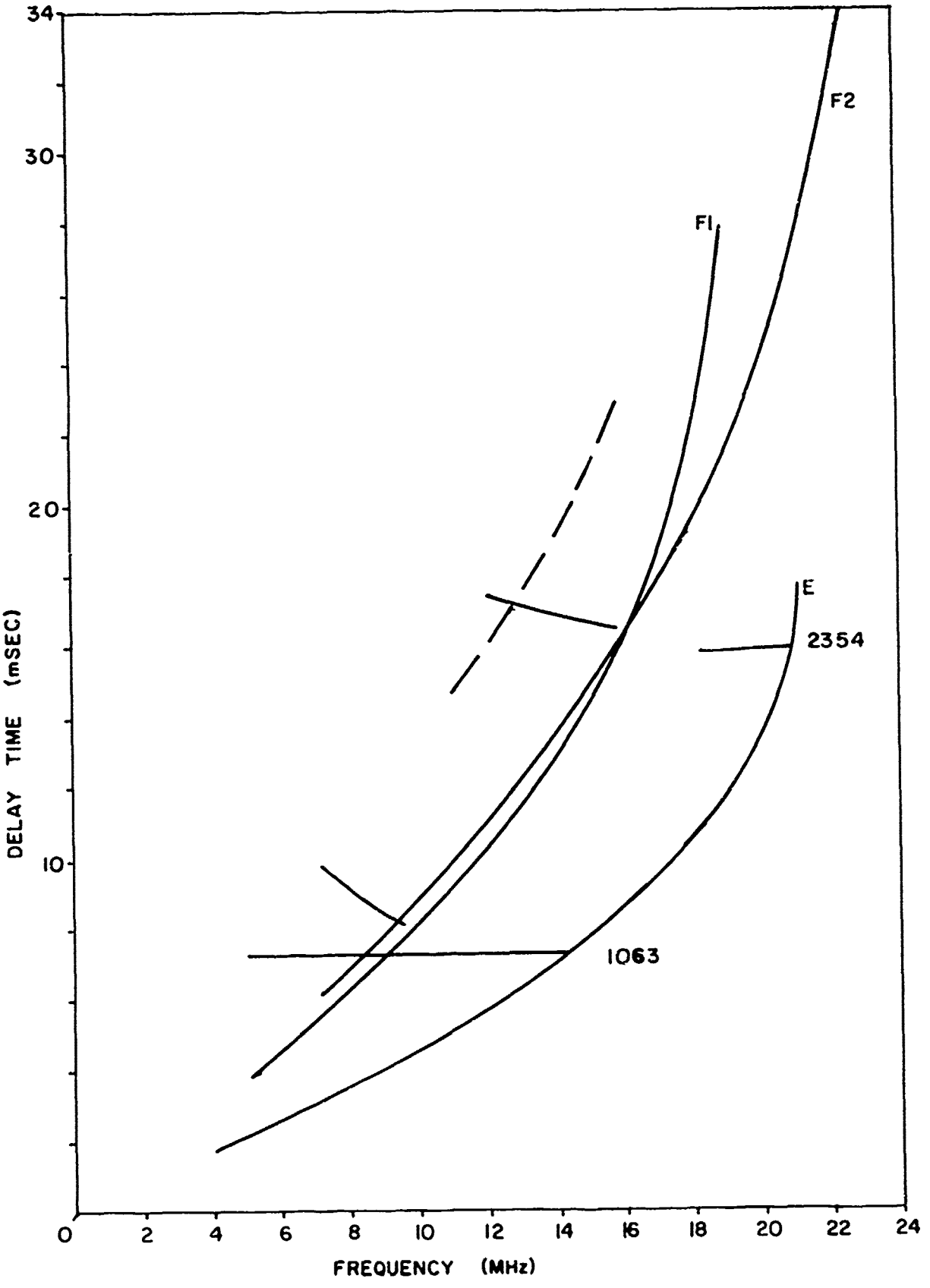


Figure 2-9. Traces of Leading-Edge-Focused Ground Clutter and Discrete Targets as Shown in Figure 2-2. A soft boundary caused by defocusing is indicated by the dashed line.

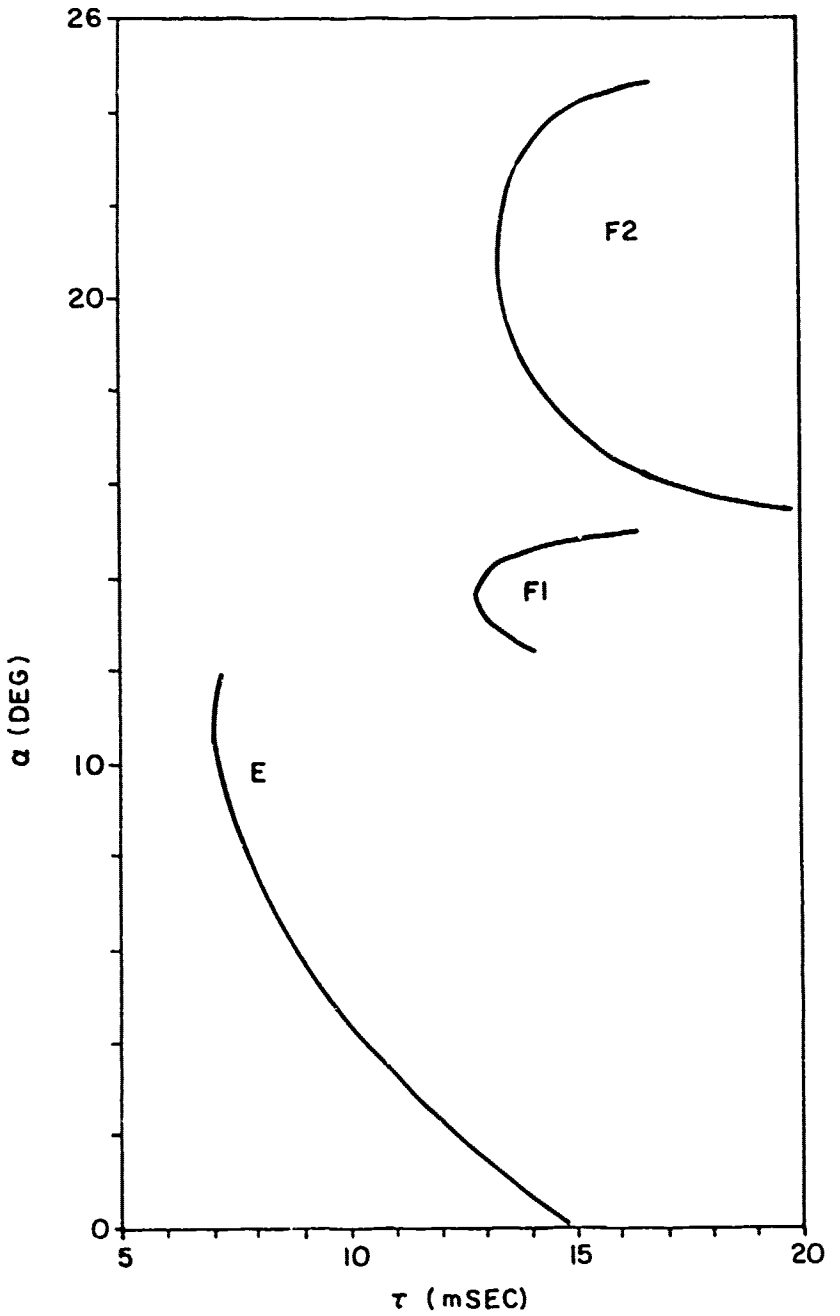


Figure 2-10. Traces in the Radar Delay-Elevation Angle Plane of Rays Representing 14-MHz Propagation via E, F_1 and F_2 Layers to the Earth's Surface. Defocusing is indicated by decreased slope, focusing by increased slope.

of underlying layers. In particular, note that the lower portions of the curves for the F1 and F2 layers enter a cusp, near 12 ms and 15.5 ms, respectively, that represents the transition to the underlying layer. As the cusp is approached, the curve approaches the horizontal. This is indicative of defocusing. Thus, the strength of ground clutter echoes corresponding to these points on the curve can be expected to be weak. For example, for the F2 layer, a strong focusing is indicated at about 13 ms, and as the range increases beyond 13 ms, the low angle rays become increasingly spread and would fade completely by the time a 20 ms range is reached. This is a good example of a soft boundary attributable to the obscuring effect of an underlying layer. The boundary is "soft" because defocusing leads to a gradual diminution of echo strength with increased range. The dashed line in Figure 2-9 indicates how this soft boundary would limit the F2 mode observable region. The precise position of this boundary would vary with those factors influencing the radar's sensitivity such as path loss, noise level, and target cross section.

In some cases, the maximum delay of a mode's observable region occurs at a hard boundary--one at which focusing occurs. The maximum range (or delay) focusing has been studied by Croft (1967) who has shown that its occurrence depends upon the distribution of electrons between two ionospheric layers. Examples of this can frequently be observed in the frequency scans of an OTH radar; Figure 2-11 is a good example of this. The time axis of the oblique ionogram is too compressed to indicate the details of the repeater-generated oblique ionograms in the vicinity of the trace of the maximum range-focused clutter. Croft has treated the matter in some detail and has found that an oblique ionogram trace will undergo an S-shaped turn in making the transition from the low ray trace of an overlying layer to, ultimately, the low ray trace of the underlying layer. The extension of the repeater trace beyond the trace of the maximum range-focused clutter is caused by low-ray propagation supported by the underlying layer, so it can be seen to occur at a slightly lesser delay.

24 October 1971
1616Z
Transmit Steer 290°
Receive Steer 288.5°

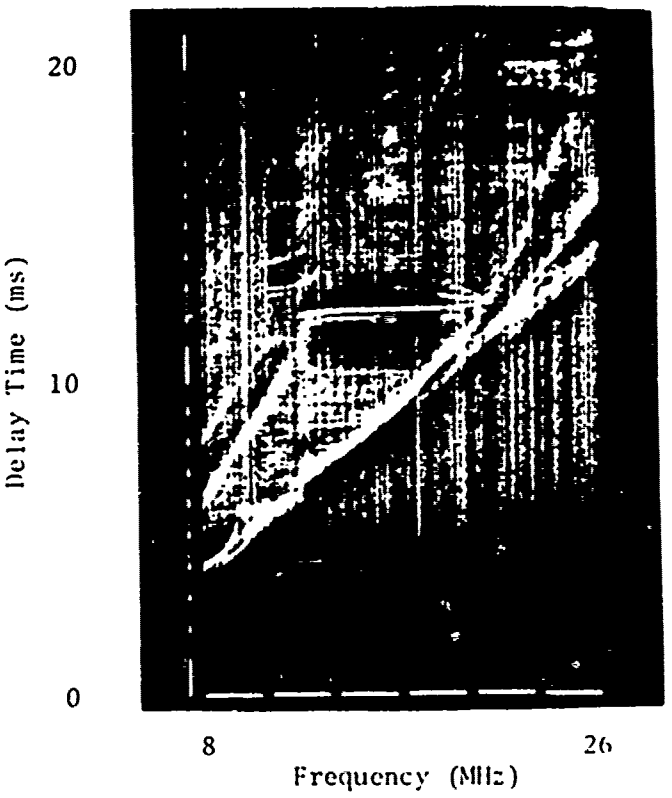


Figure 2-11. Skywave Backscatter f-Scan Showing a Hard MOR Boundary Caused by Maximum Range Focusing

From the above discussion, we may summarize the following points:

(1) The leading edge (in radar time delay) of a mode-observable region is usually delineated by the hard boundary produced by the familiar focussing effect.

(2) The trailing edge of a mode-observable region may be either a "hard" or a "soft" boundary, depending upon the structure of the layer and upon other contingencies. A soft boundary's position will vary with path loss, noise level, and the radar cross section of the particular target. Consequently, its position can be predicted and used in mode identification decisions only with careful consideration of statistical detection theory.

2.2.2 Occurrence and Extent of Ambiguous Regions

In this section, some examples of the extent of ambiguous regions in the frequency-radar delay plane are given. A comprehensive analysis of the occurrence of ambiguities is not intended; only a demonstration that the problem occurs a significant fraction of the time.

Nighttime conditions at temperate latitudes represent the simplest category of propagation with which the OTH radarist must contend. The underlying F1 and E layers are controlled by sunlight, and their effect disappears shortly after sunset. The F2 layer, although also produced by insolation, responds with a significant time lag (on the order of four hours). Consequently, residual ionization prevails through much of the night, so the ionosphere has the form of a single layer. Figure 2-12 is an example of a sky wave backscatter trace obtained under these simple layer conditions. Note that the trace of repeater returns at 12 ms extends continuously from the minimum frequency used by the radar to the boundary of the mode-observable region delineated by the minimum-range focussed clutter. Were an operating frequency near 13 MHz used, the radar would cover an interval of radar delays extending from about 8 ms to greater than 17 ms. The latter figure would be a soft boundary and cannot be precisely specified. Figure 2-13 shows a similar trace with improved time resolution. Note that although the pure high-ray trace is not apparent, there appears a trace caused by mixed high-ray and low-ray propagation. The mode-observable region for this mixed-mode extends from the hard boundary at the trace of the minimum-range-

25 April 1971
0205 Z

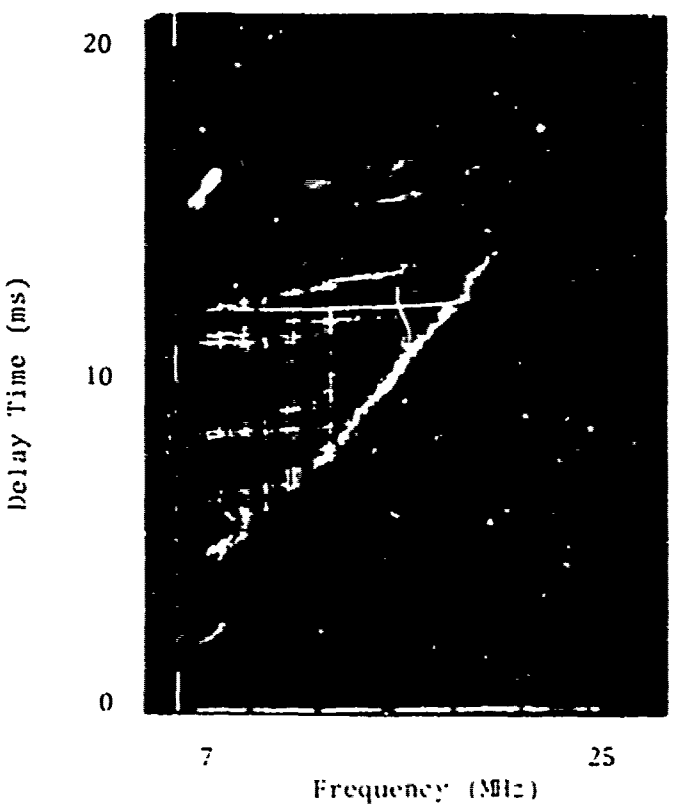


Figure 2-12. Skywave Backscatter Trace Showing a Single-Layer Ionosphere, Typical of Winter Nighttime Conditions

25 April 1971
0229 Z

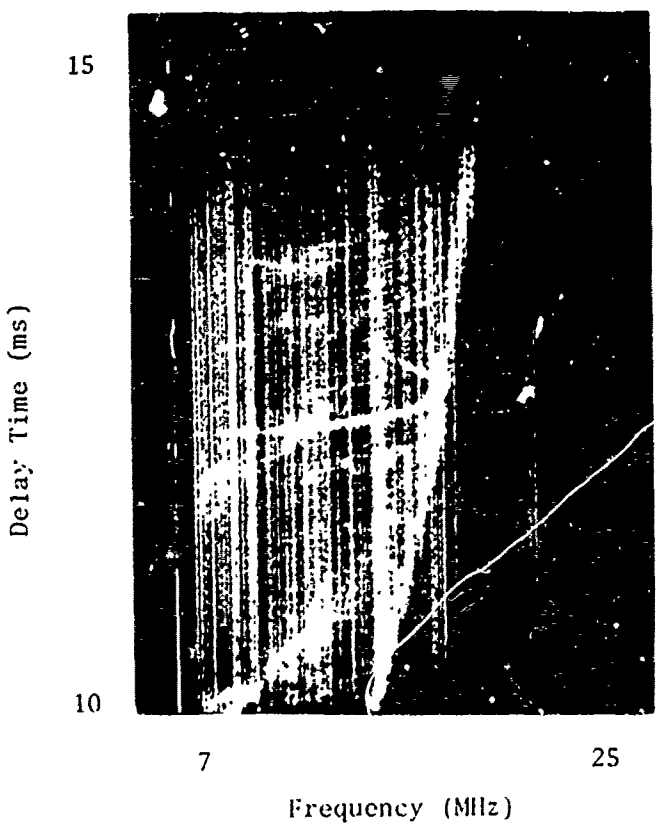


Figure 2-13. High Resolution Skywave Backscatter f-Scan Made when the Ionosphere Consisted of a Single Layer. A region of ambiguity occurs because of high ray-low ray mixed propagation within 3 MHz of the leading edge hard boundary.

focused clutter which it shares with the mode observable region (MOR) of the pure low ray mode, to a "soft" boundary which for the repeater signal lies approximately 3 MHz below the frequency of the hard boundary. Measured on the radar delay axis, the soft boundary of the high-low ray MOR lies about 2.5 ms beyond the "hard" minimum delay boundary. This high-low ray MOR constitutes an ambiguous region because it is a subregion of the low-low ray MOR. Consequently, if as in the example mentioned earlier, a 13-MHz operating frequency is used for the radar, the interval from 8 to 10.5 ms would be ambiguous, and the region 10.5 to 12 ms or more would not be. However, the trailing boundary of the ambiguous region is soft, so for targets with cross sections smaller than that simulated by the repeater, the ambiguous region would be less extensive. Figure 2-14 shows a similar repeater-generated oblique sounding record over the same Whitehouse-to-Lincoln path, but this time for late afternoon conditions. Again, the high-low ray mode extends about 3 MHz below the hard boundary formed by focusing at the mode's leading edge. The extent of the ambiguous region is about 3 ms in this case. There is some indication of a high-high ray mode trace, but its strength is significantly less than that of the other modes.

A set of measurements by Tornatore (1972) indicates that the low-high mode would be weaker than the low-low mode by about 3 db, one MHz below the maximum frequency limit, and about 6 db at 2 MHz below the hard boundary (MUF). Thus, a radar sensitive enough to have a 3-db margin of detection for targets of interest would have an ambiguous region extending about 1 MHz below the maximum frequency, or a radar delay of about one ms beyond the near-range focused clutter trace. This kind of ambiguity can almost always be expected even with the simplest ionospheric structure.

Also, multimoding and ambiguity can occur under single layer conditions because of multiple-hop modes. Figure 2-15 is an example of such a skywave backscatter record. A repeater-generated trace produced by single mode propagation occurs at 12.2 ms radar delay and joins with the high-low ray trace at the focused clutter trace. A second repeater trace, considerably

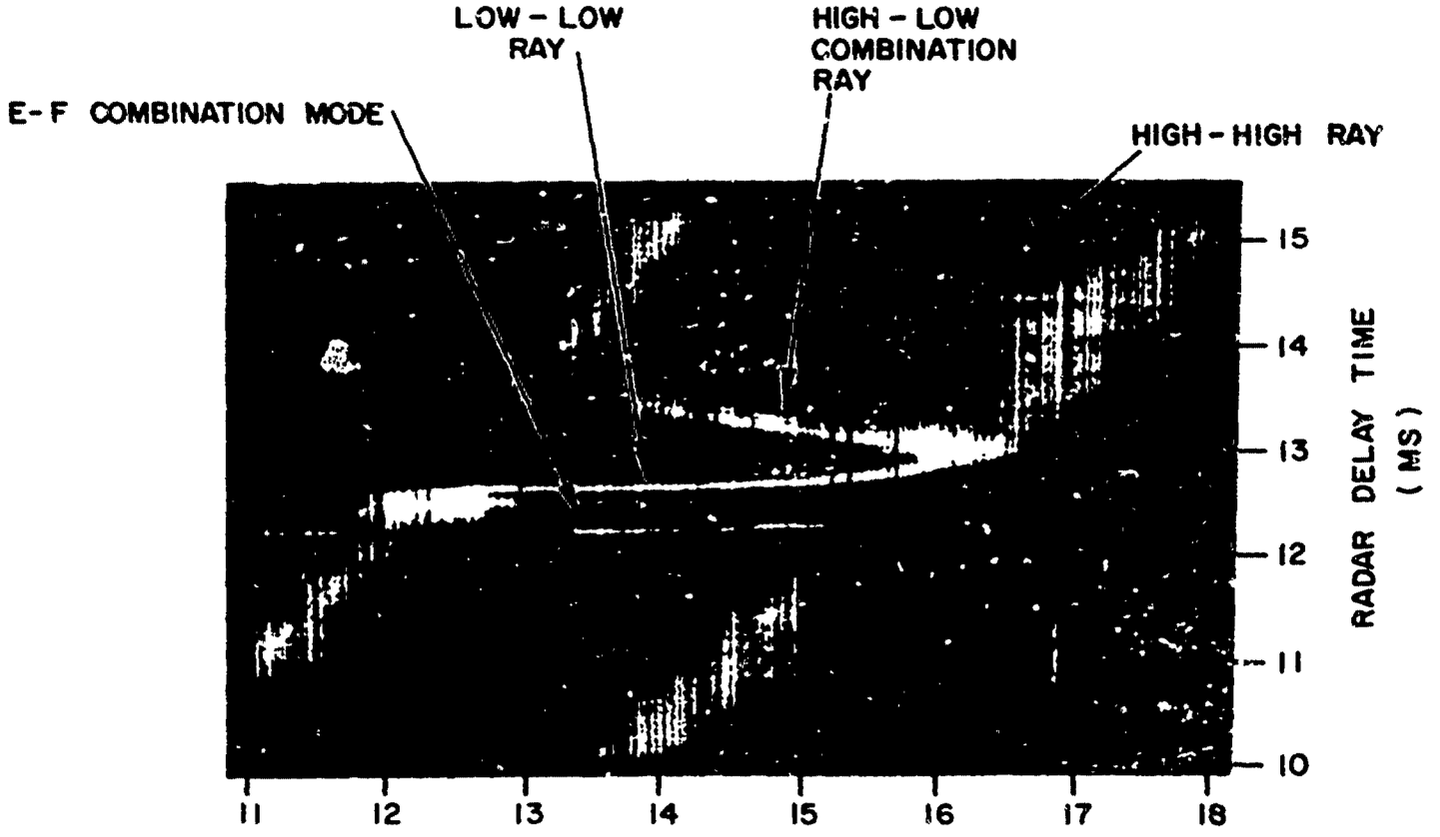


Figure 2-14. A Typical Example of an Oblique-Ionogram Showing Echoes Over Various F-Mode Propagation Paths to Lincoln, Nebraska.

24 Octob · 71
2122 Z
Transmit Steer - 290°
Receiv Steer - 288.5°

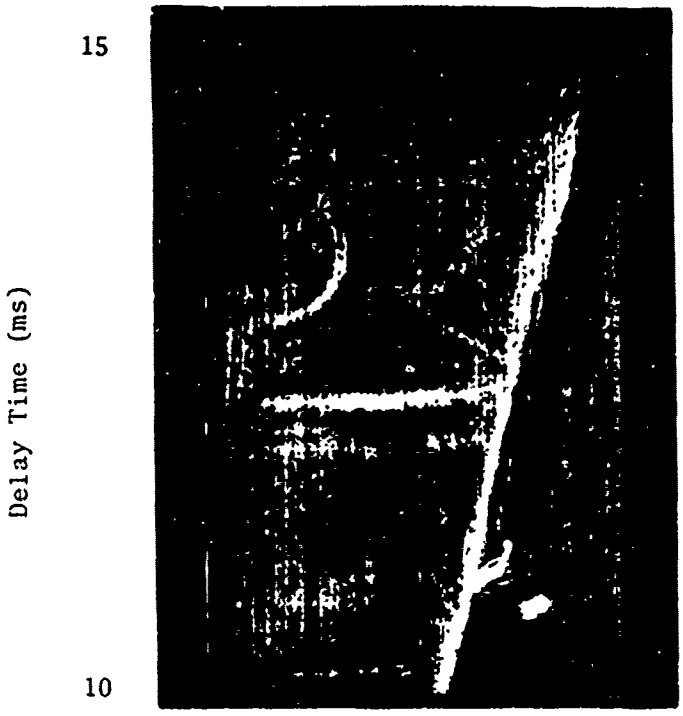


Figure 2-15. Skywave Backscatter f-Scan in Which an Ambiguity Occurs Because of Multiple-Hop Propagation

dispersed, appears at ranges from 12.9 to 13.7 ms. Barely perceptible is yet a third trace extending in radar delay from 13.5 to beyond 14 ms. This latter trace results from propagation to and from the repeater via a two-hop mode, while the stronger trace is the result of a mixed mode (e.g., propagation to the target via a one-hop mode and return via a two-hop mode). In this case the observable region for the mixed one-hop, two-hop mode has a high frequency boundary that is about 9 MHz below that of the one-hop mode. This difference in frequency will vary with the range of the target (or repeater) being considered and with the details of the layer's electron density distribution

For normal, effective heights of the F layer (under 400 km), the maximum frequency at which two-hop propagation is supported to any given range is usually less than 75% of the frequency at which one-hop propagation is supported. Under nighttime conditions, when the F layer critical frequency can be as low as 3 MHz, the 4000 km-range MUF can be as small as 10 to 12 MHz, resulting in a frequency interval between the MUF of the one-hop and two-hop modes of approximately 3 MHz. The difference in frequency between the leading boundary of the one-hop MOR and the mixed one-hop, two-hop MOR would be nearly as great, perhaps about 2.5 MHz. It is conceivable that under these conditions a well developed one-hop high-ray trace in combination with such a minimal frequency difference between the leading boundaries of the one-hop and two-hop mixed mode MOR's could preclude the possibility of choosing an operating frequency at which the radar delay of interest would be free of ambiguity; however, a brief search through available skywave backscatter records failed to discover any such examples. A number of examples were found in which a region of one-hop, low-ray F-mode propagation occurred unambiguously over a 3-MHz frequency interval.

To summarize this discussion of the simple, single-layer ionosphere:

(1) Such conditions occur frequently at night.

(2) The one-hop, low ray MOR (MOR-1FL) extends from the frequencies of 7 MHz or less to the hard boundary delineated by the focused clutter trace.

(3) MOR-1FL is ambiguous with the one-hop high-ray MOR (MOR-1FH) in a region extending 1 to 2 MHz below the one-hop F focussed clutter trace. The low frequency boundary of the ambiguous region is soft.

(4) The MOR-1FL is ambiguous with the mixed one-hop, two-hop MOR (MOR-1F2F) at frequencies less than 60% to 75% of the 1FL mode MOF for any range.

(5) The high frequency boundary of the MOR-1F2F is a hard boundary, but it is usually not discernible in the skywave backscatter records as a trace of focused clutter.

Figure 2-16 diagrams the MOR's discussed above. At least a short interval of radar delay will usually be free of ambiguities when such conditions prevail.

The next, more complicated, circumstance is that in which the F layer is accompanied by an underlying layer. This may be either the E layer at 100 km or the F1 layer, which occurs near 200 km. Of course, both of these underlying layers exist simultaneously, but for the moment we will consider the effects of only one.

A sufficiently dense underlying layer could entirely obscure the F layer, so no trace indicative of the F layer would appear on the record. This does not happen normally but may happen in cases of intense sporadic E. A more common occurrence is for the underlying layer to obscure all but those backscatter echoes that occur in the vicinity of the leading focused clutter trace. An example of this is shown in Figure 2-17. An expanded trace that better shows the limiting of the backscattered echoes to within a MHz of the leading edge is provided. As noted earlier, the region of the one-hop F MOR within a MHz of the leading edge is subject to ambiguity involving the high and low rays, so this represents a case where potential ambiguity is unavoidable. The portion of the E region MOR at frequencies higher than the F layer's focused clutter trace is unambiguous and could be used by the radarist. However, the ranges of the scatterers observed via this mode are substantially less than those observed via the F mode. The low frequency boundary of the E region MOR is a soft one caused by absorption of the propagated waves in the D layer. Figure 2-18 is a schematic diagram showing the distribution of the mode-observable regions in the radar delay-frequency plane. The high frequency edge of the F layer MOR is shown as a hard boundary, but the oblique ionogram indicates a raggedness that makes the position of the

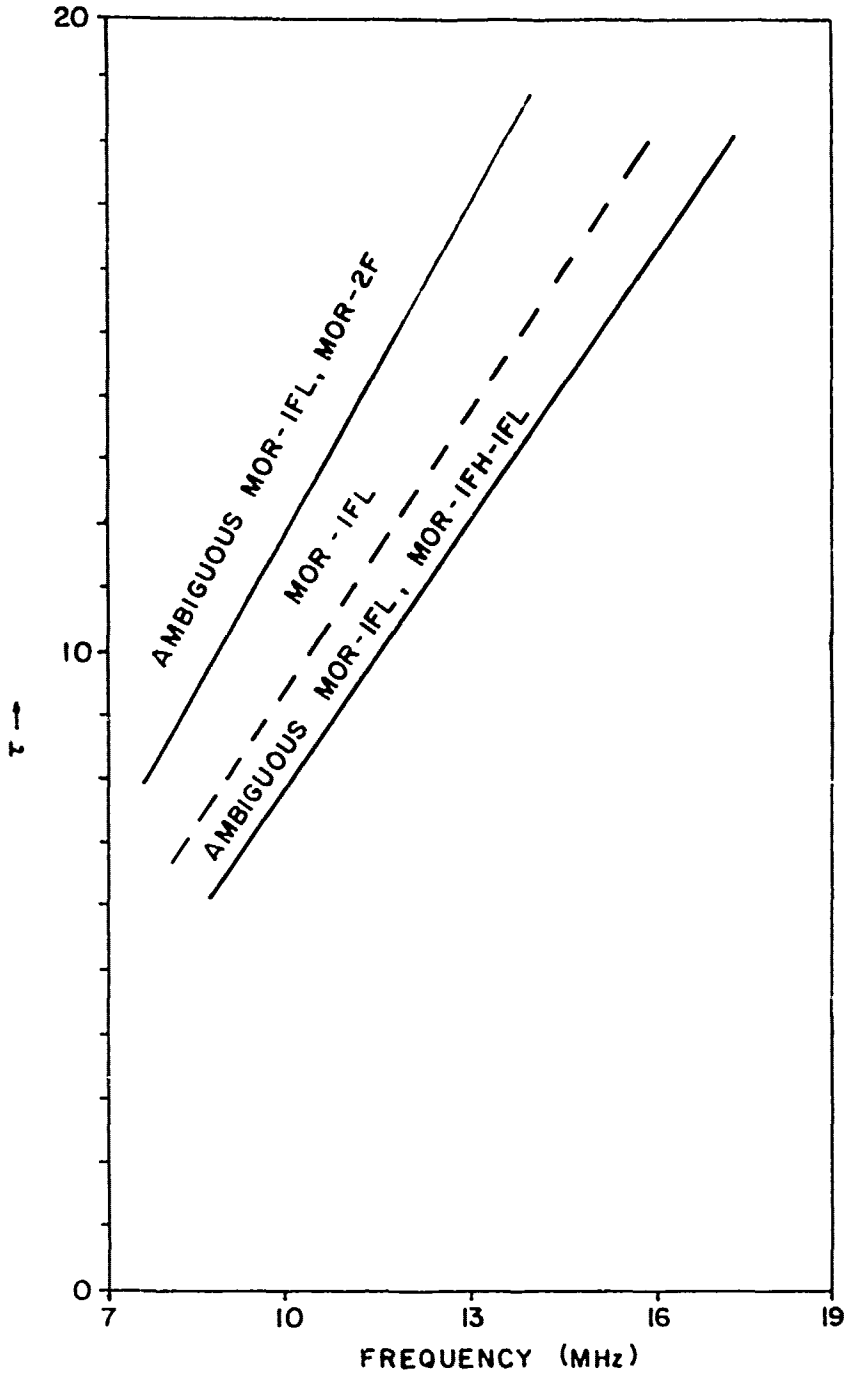


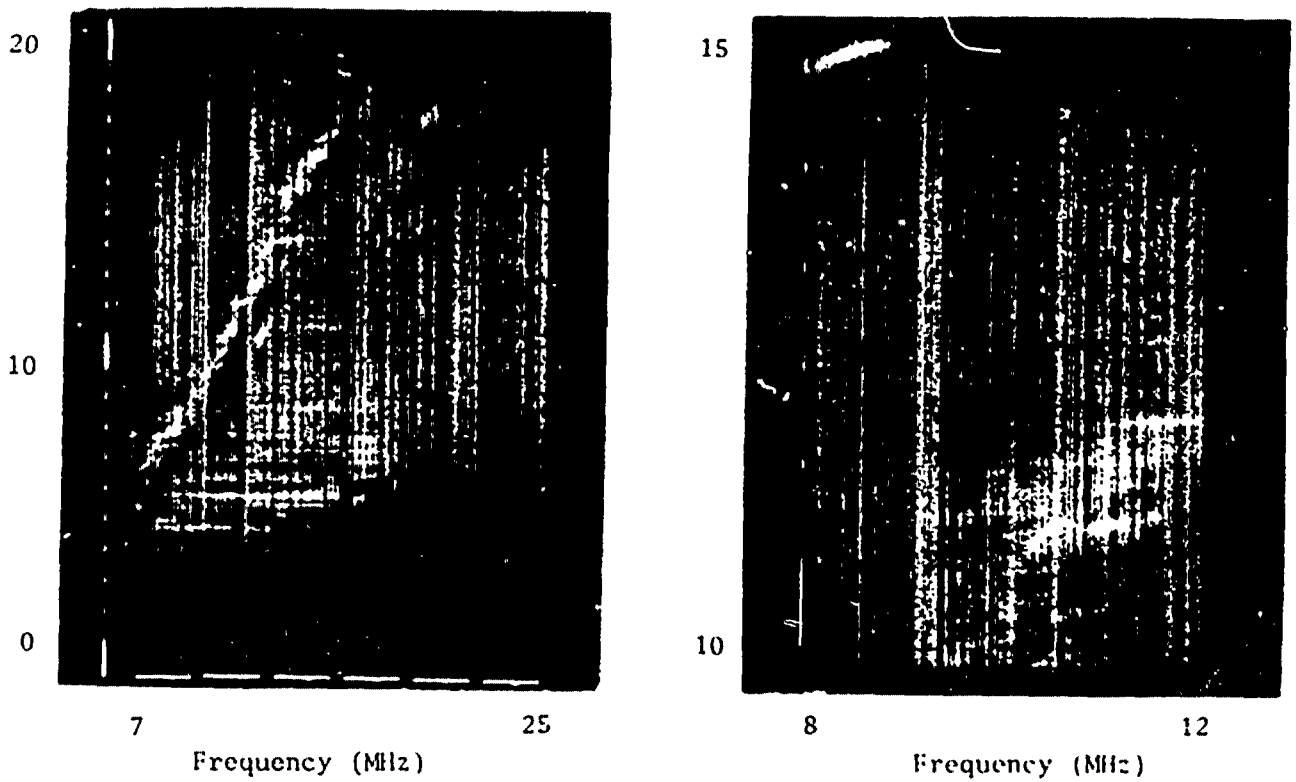
Figure 2-16. Single Layer Ambiguities (Nighttime Conditions)

A75-175

12 August 1971
0066 Z

12 August 1971
0010 Z

Transmit Steer- 290°
Receive Steer - 288.5°



(a) (b)

Figure 2-17. Skywave Backscatter f-Scan Showing How Echoes can be Restricted to the Immediate Vicinity of the Leading Edge of an MOR. Figure 2-17(b) is a high-resolution presentation of a portion of Figure 2-17(a).

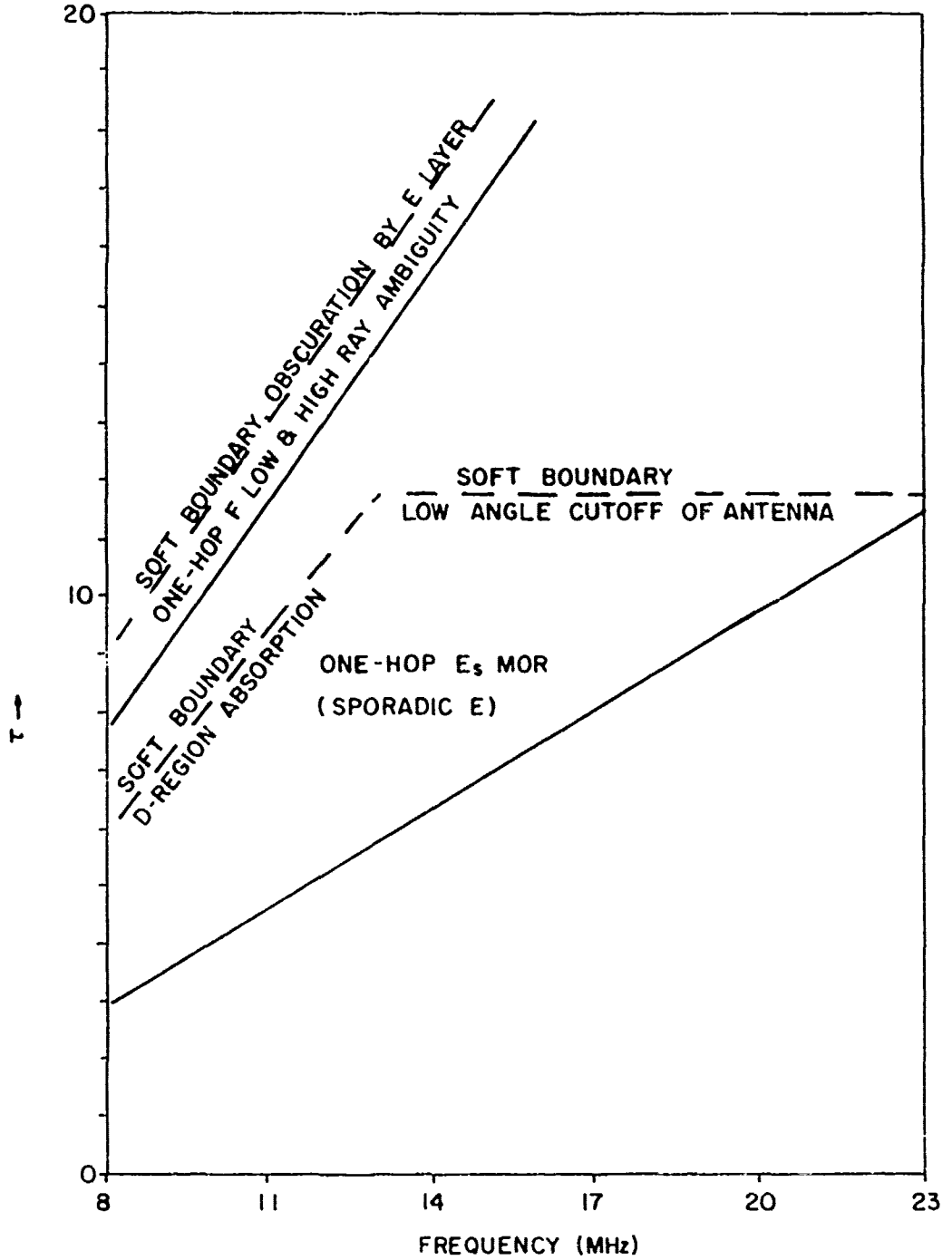


Figure 2-18. Schematic Diagram Showing how the MOR of a Higher Layer can be Limited by the Obscuration of a Lower Layer

A75-176

boundary uncer• n. This may be the effect of irregular obscuration by the underlying sporadic E and to various scattering effects caused by the same layer.

In the preceding aiscussion, the underlying layer had an MOR that extended to higher frequencies than did the F layer MOR. We next consider two cases where the F layers MOR extends to frequencies higher than does the MOR of the underlying layer. These cases were discussed in the earlier section on MOR boundaries and to much greater length by Croft (1967). Here, we simply note that one or the other of these cases will occur during winter daytime conditions, when the F layer critical frequency is anomalously high, and the E layer critical frequency is somewhat reduced by the sun's greater zenith angle. Figure 2-11 showed an example in which the low frequency edge of MOR-1FL occurs at a focused clutter trace, while the example in Figure 2- 19 is of a MOR-1FL that is softly bounded, so the repeater trace fades away at its low frequency limit.

In either case the frequency extent of the MOR-1FL increases with radar delay, being about 4 MHz at a radar delay of 7 ms and increasing to 8 MHz at 12 ms in Figures 2-11 and 2-19 . Also, in each example, the low frequency portion of the MOR-1FL intersects the MOR-1E and MOR-1E,1F, indicating a region of mode ambiguity. The ambiguous region can be barely discerned as a region of enhanced clutter in the radar delay interval from 6 to 10 ms. A more clear delineation of the ambiguous region appears in Figure 2-20 where there is some minimum range focusing of the clutter seen via the 1E mode.

Figure 2-21 is a schematic diagram of the distribution of mode observable regions under these conditions. Note that there is considerable opportunity for ambiguity in the vicinity of the low frequency edge of the MOR-1F, where it overlaps the MOR's of the E layer and is near the MOR of the two-hop F mode, as well as the MOR's of the mixed modes. However, it may also be noted that there is an interval of several MHz in which the MOR of the one-hop F layer low ray is unambiguous.

4 November 1971
1342 Z
Transmit Steer 290°
Receive Steer 288.5°

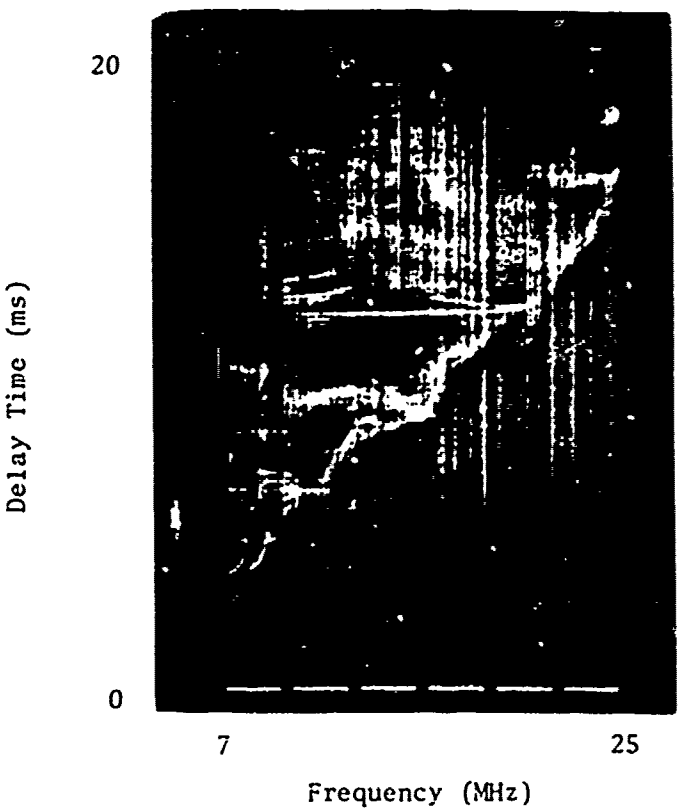


Figure 2-19. Skywave Backscatter f-Scan Illustrating a Softly Bounded Low Frequency Limit to the Mode Observable Regions

24 October 1971
1959 Z
Transmit Steer - 290°
Receive Steer - 288.5°

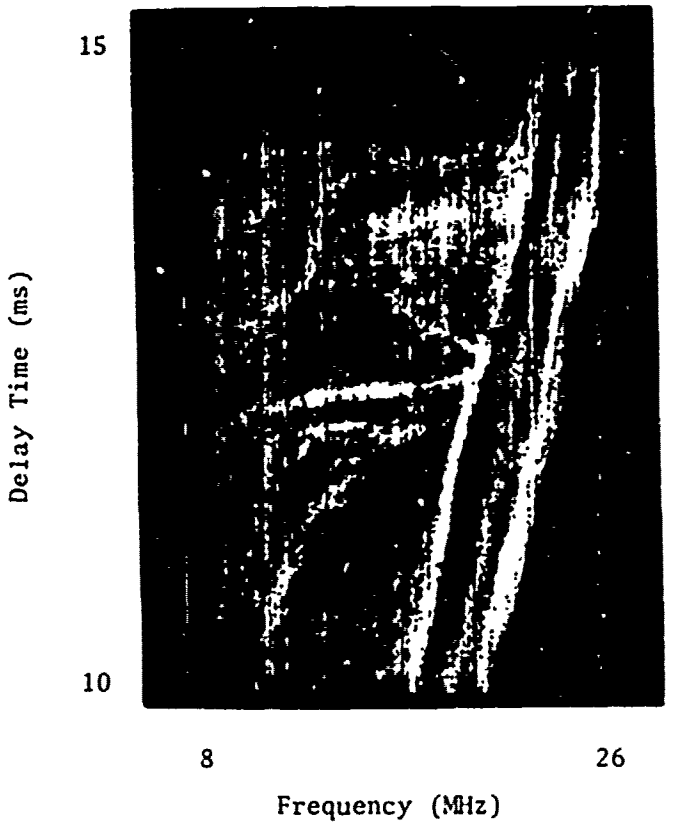


Figure 2-20. Skywave Backscatter f-Scan Illustrating an Ambiguous Region Where MOR-1FL Intersects MOR-1E and MOR-1E,1F. Minimum range focusing of the 1E mode is apparent.

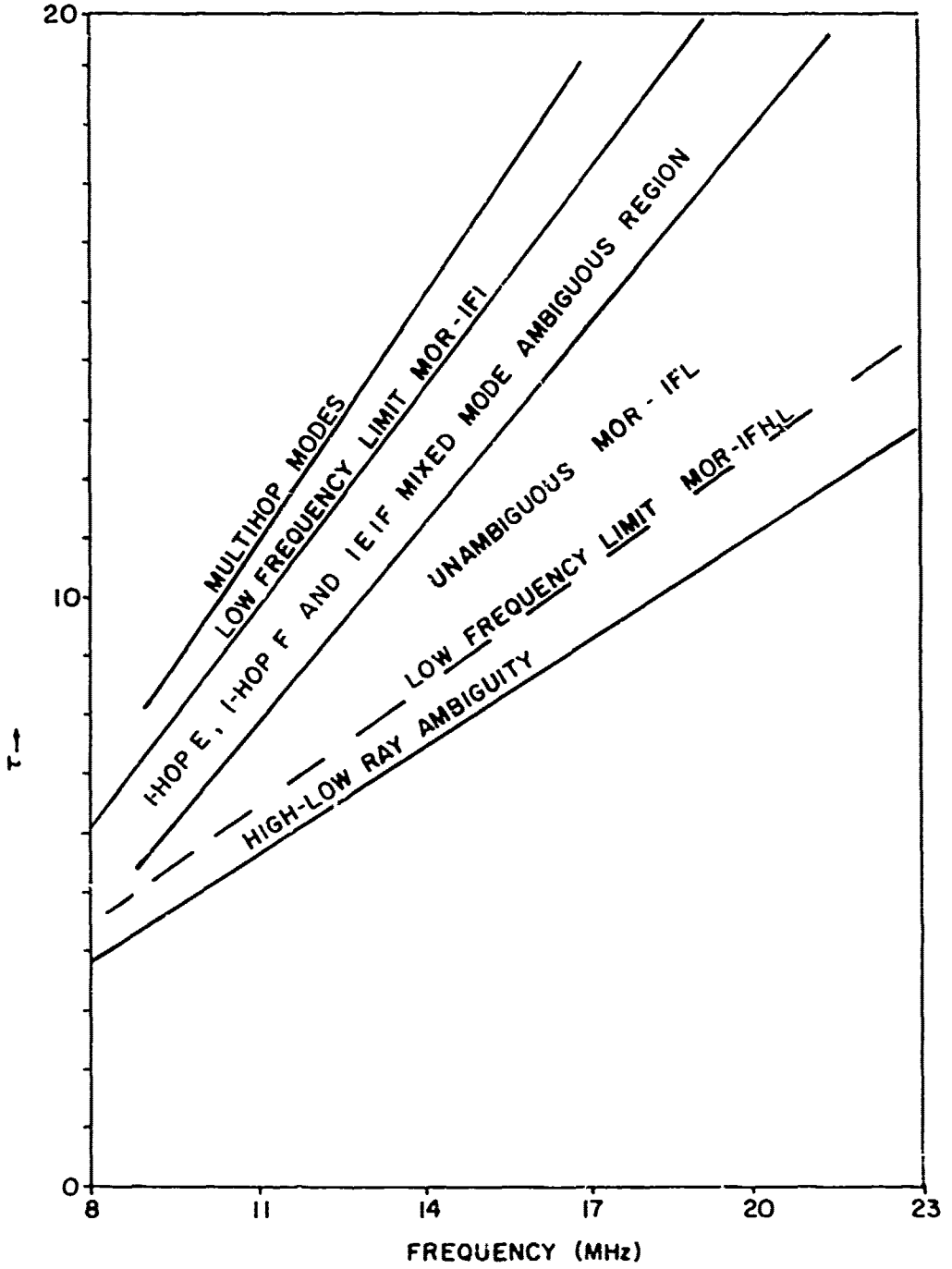


Figure 2-21. Schematic Diagram of Mode Observable Regions which Exist when the Higher (F) Layer's MOR Extends to Higher Frequencies than does the Underlying Layer's

A75-177

The three classes of ionosphere discussed above include an important fraction of the situations with which an OTH radar must deal, respectively representing nighttime conditions, summer daytime and winter daytime. An experienced HF radarist will recognize that the simple categorization made here does not always hold. For example, during summer nights the sporadic E layer may introduce additional modes in the manner of the daytime E layer, and during ionospheric storms the F layer critical frequencies may be reduced by 15% to several tens of percent (Matsushita, 1959) depending upon the geomagnetic latitude of the reflection point. Also, there are the transitions between daytime and nighttime conditions. During what we have here proferred as typical winter conditions, no serious problem arises because the MOR-1FL for both daytime and nighttime conditions extends to higher frequencies than do those of the underlying layers. However, during summer conditions, a transition must take place wherein the high frequency boundary of the E layer MOR must move across that of the F layer. At ranges beyond those at which one-hop E can occur, there is still no problem. The only effect is the one of obscuration, which does not occur until the E layer MOR at any given radar delay extends to appreciably higher frequencies than does the F layer's. However, at E layer ranges (out to about 1200 km or about 10 m. radar delay) a well developed ambiguity occurs. Figure 2-22 diagrams the MOR's that occur during such a transition. During the hour following 1800, the E layer MUF drops rapidly and by 1900 lies a few MHz lower in frequency than the F-layer MUF.

This completes the discussion of the occurrence and extent of mode ambiguity. We have noted that, provided one is free to select the radar's operating frequency solely on the criterion of nonambiguity of the radar echoes received, there usually can be found, for any radar delay, a few MHz interval free from ambiguities. An exception to this sometimes occurs during summer daytime conditions when obscuration by underlying layers may limit the F layer MOR to a band extending only a MHz or so below the focussed minimum-range clutter trace. Under such conditions, propagation utilizing the high-ray mode on the outward or inward path produces a delayed trace that has a signal strength on the order of 3 db less than that of the echo propagated entirely by the low ray mode.

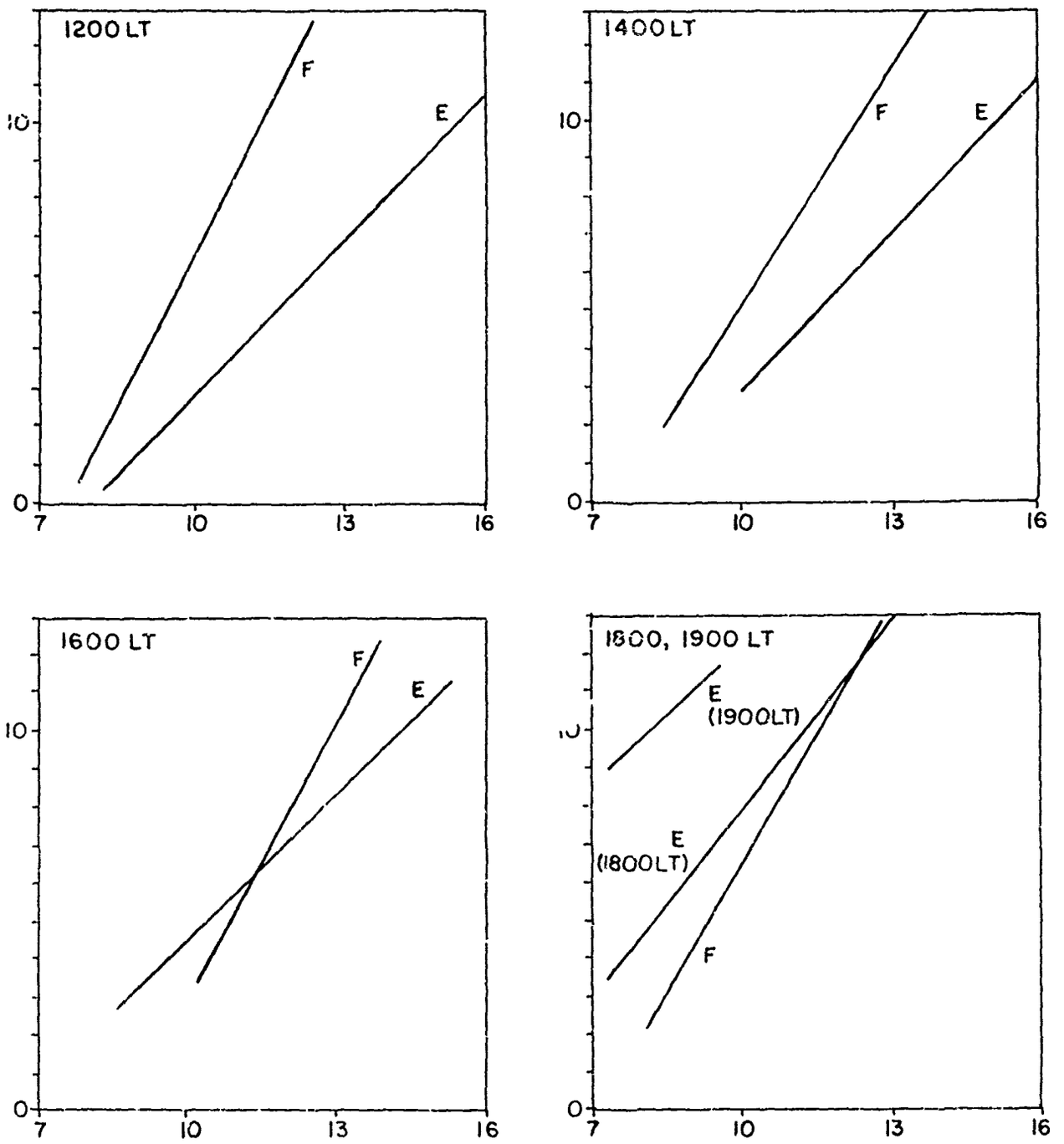


Figure 2-22. Diagram Illustrating the Transition from Summer Daytime Propagation Conditions (1200 LT), when MOR-1E Extends to Higher Frequencies than does MOR-1F, to Nighttime Conditions (1900 LT), when the Converse is True

A75-178

Mode ambiguity occurs also at short ranges at which E layer propagation can occur. A particular problem is posed by periods of day/night transition, when the leading edge traces cross.

2.3 References

1. Croft, T.A., "HF Radio Focusing Caused by the Electron Distribution between Ionospheric Layers," J. Geophys. Res., 72, 9, 1 May 1967, pp. 2343-2355.

2. Matsushita, S., "A Study of the Morphology of Ionospheric Storms," J. Geophys. Res., 64, 3, March 1959, pp. 305-321.

3. Tornatore, H.G., "The Relative Amplitude of Low- and High-Ray-Propagated F-Mode Signals," Project Report No. 205, ITT Electro-Physics Laboratories, Inc., Columbia, Maryland, January 1972.

3. TESTS FOR MODE IDENTIFICATION INVOLVING BOUNDARIES

In this section it is assumed that individual detections by an OTH radar are first processed and assembled into tracks. In the present context, a track is a subset of the detections, which, on the basis of proximity to one another and consistency of doppler shift, range rate, and direction, appear to be the results of multiple observations of a single target. If the tracking is done in the space of radar observables (i.e., frequency, radar delay, and antenna angle), a single target that is illuminated by multiple modes may give rise to multiple tracks. If, on the other hand, the tracking is to be done in the form of geographic (or other true space) coordinates, it is necessary to perform coordinate conversion on a detection-by-detection basis. In this case, if a detection happens to lie within an ambiguous region of the τ-f plane, multiple hypotheses would have to be carried regarding the operative mode of propagation, resulting in multiple hypothetical positions for the detected target, and consequently, multiple tracks. Thus, in either event the effect of multiple propagation modes is to raise the possibility of there being more radar tracks than there are targets. For the present discussion, we assume that the tracking is to be done in the τ-f plane. Also we begin with the discussion of the more general case of a frequency-hopping radar. The problem of mode identification for a fixed-frequency radar is covered in later portions of this report.

3.1 Tracks in the f-τ Plane

If a fixed target is observed by a frequency-hopping radar, the individual observations will lie on a trace that is equivalent to an obliquely sounded ionogram made over the path. The detections of an approaching (or receding) target would form a cloud of points that would be bounded by the oblique ionogram traces for paths to the near and far (with respect to the radar) points of the target during the period of observation. (See Figure 3-1.)

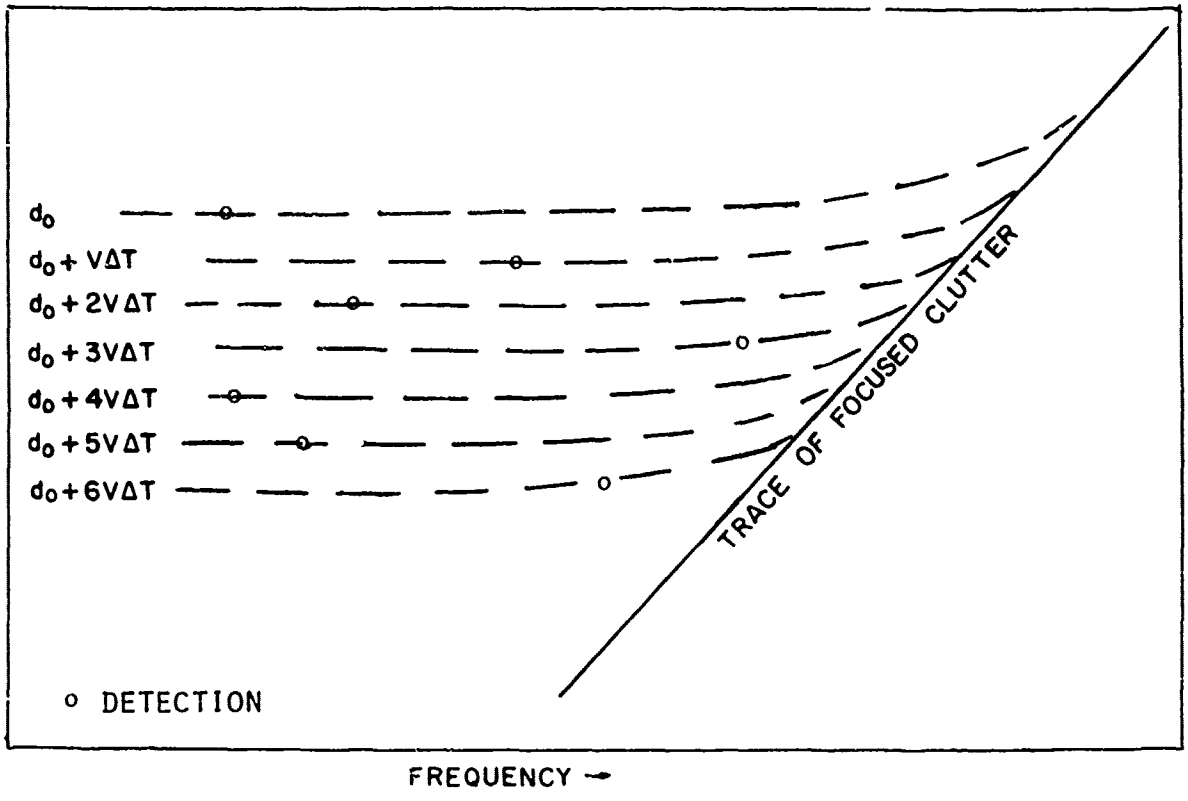


Figure 3-1. Occurrence of Detections in the τ-f Plane When an Approaching Target is Observed with a Frequency-Hopping Radar

If the radar's range resolution (bandwidth) is not great and if the period of observation is sufficiently small, the displacement during the course of the observations may be small compared to the radar's resolution, and the observations may all fall on a single oblique ionogram trace (although given sufficiently great signal-to-noise ratio, the displacements could be measured). Assuming 10-second dwells and a 900 km/hr. aircraft, the displacement would be about 2.5 km during a dwell and would be discernible from one observation to the next, only if the radar's bandwidth exceeded 60 kHz. With a 20-kHz radar, an observation at any time would have a radar delay discernibly different from those made more than two dwells earlier or later, while a radar with a 250-kHz bandwidth would see an appreciable change in range from one dwell to the next. Also, a wide bandwidth system (≈250 kHz) would experience the effects of dispersion in the ionosphere so that observations of a target made at frequencies near the MUF would be at discernibly greater radar delays than would those made well below the MUF.

The above properties are of importance to the tracker, which must take account of them either by correct prediction of, or by being tolerant of, the effects.

3.2 Tests for Mode Identity

The mode identification tests to be discussed in this section all involve the behavior of the signatures at the boundaries of the mode observable regions. The basis for these tests is that a track must be entirely confined within the mode observable region (MOR) for the operative mode. If a track includes points outside of a particular MOR, then one may conclude that the corresponding mode was not the mode that provided the detections which made up the track. Conversely, if all of the detections occurred within a particular MOR, then one may be justified in identifying the propagation mode of the detections as that corresponding to the MOR. However, now there must be a certain amount of reservation in making the identification. For example, the set of detections may also all be within a second MOR as well; i.e., the

detections may all lie within an ambiguous region. In order to resolve such an ambiguity, observations must be made such that a detection occurs (or is observed not to occur) within an unambiguous region, or within some combination of ambiguous regions which provides an unambiguous result.

Another concept is that the credibility of an unambiguous detection, or of an unambiguously missed detection, must be considered. An OTH radar often operates in a noisy environment and is subject to false alarming. Also, the signals fade because of various propagation effects, so missed detections are to be expected. The relative probability of these two kinds of error are determined by the thresholds employed in making detection decisions. Clearly, a few detections outside of an MOR should not rule out the associated mode if the likelihood of such detections occurring erroneously is large. Similarly, the absence of detections outside of an MOR cannot be regarded as a conclusive indication of mode if there is an appreciable probability that the detections could have been missed by chance. Thus, mode identification by boundary crossing involves considerations of probabilities, even in the idealized case where the boundaries are well defined and predictable. At soft boundaries, where the probability of detection diminishes gradually with frequency or radar delay, the need for probabilistic arguments is all the more acute.

Having discussed some of the general principles involved in this method of mode identification, we now consider some particular examples. The first example will involve the overlapping of two mode observable regions to form an ambiguous region. Figure 3-2 indicates how the MOR's might be related. Two traces of target detections are shown. The upper trace extends beyond the high frequency boundary of the MOR of Mode B. Because these detections cannot have occurred via Mode B, it may be concluded that Mode A was the operative mode. On the other hand, the lower trace terminates at the boundary. Had these detections occurred via Mode A, then they would not be expected to be influenced by the Mode B boundary. Therefore, we may conclude that

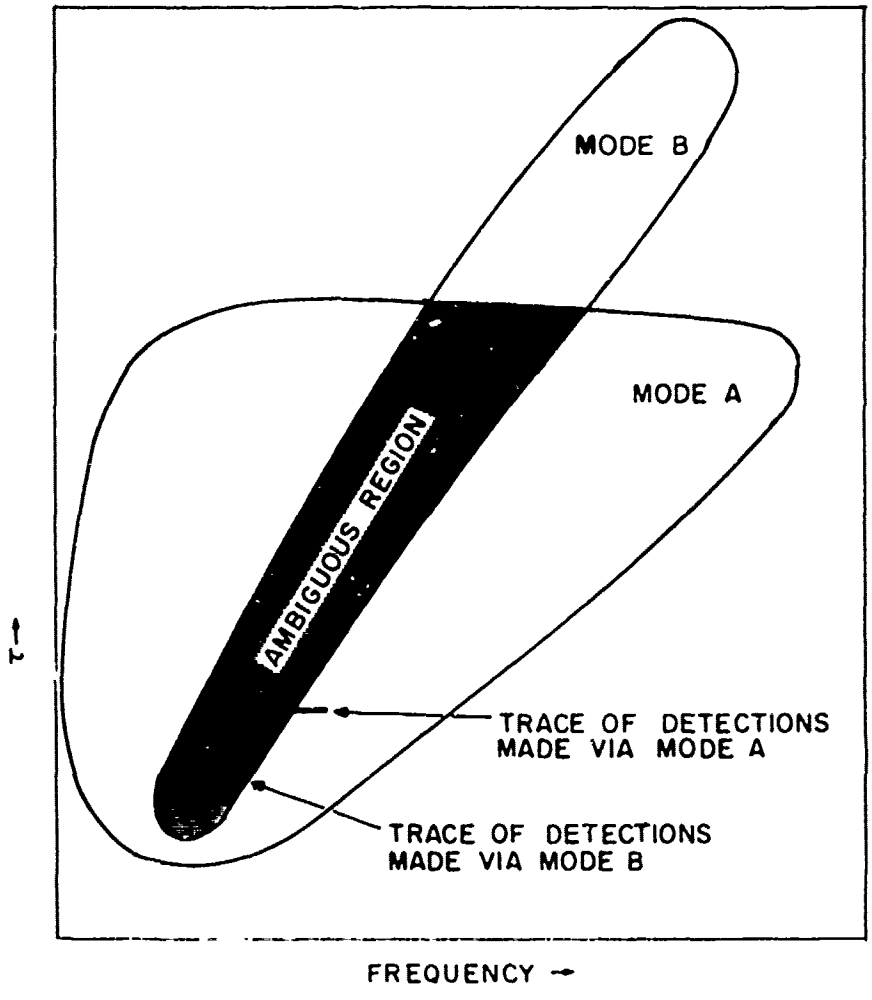


Figure 3-2. Region of Ambiguous Detections Procured by the Overlapping of Two Mode Observable Regions.

these detections were made via Mode B. However, we again emphasize that the occurrence and non-occurrence of the detections is, to some degree, fortuitous, and probabilistic considerations must be taken to ascertain a trace's extension into the unambiguous portion of the A mode's MOR.

As the next example, consider an ambiguous region produced by the overlapping of three MOR's. In principle, this could occur whenever two propagation modes illuminate a target or region of the earth's surface because a mixed (out via A, back via B) propagation mode would occur along with the two unmixed modes. Figure 3-3 illustrates the MOR's overlapping in this way. Seven traces are shown. Traces 1, 2, and 3 each cross two boundaries and can all be identified readily. Trace 1 and Trace 3 extend into unambiguous portions of the mode-C MOR, and so they can be identified as being the result of propagation via mode C. Trace 2 extends on the left into a region in which B-mode detections cannot occur and on the right into a region in which A-mode detections cannot occur. Therefore, we may conclude that the detection was effected via the C-mode. In somewhat similar fashion, truncation of the traces at a boundary would have permitted other, equally definitive conclusions to be drawn. Thus, had Trace 1 terminated at the left (low frequency) boundary of the A MOR, so that it did not extend into the unambiguous region of the C MOR, we would have concluded that A was the operative mode. Similar reasoning would have applied to Traces 2 and 3 had they been observed to terminate at an MOR boundary rather than extending beyond it.

Traces 4, 5, 6, and 7 are examples of traces which cross a single boundary. Traces 4 and 7 extend into unambiguous subregions of the C MOR and so must have been effected by the C mode. However, the ambiguity of Traces 5 and 6 cannot be resolved. Each of these traces extends from a region of three-fold ambiguity into a region of two-fold ambiguity, so the two-fold ambiguity remains. Thus, for example, the extension of Trace 6 beyond the ABC ambiguous region into the BC ambiguous region leaves the choice between B-mode and C-mode propagation unresolved.

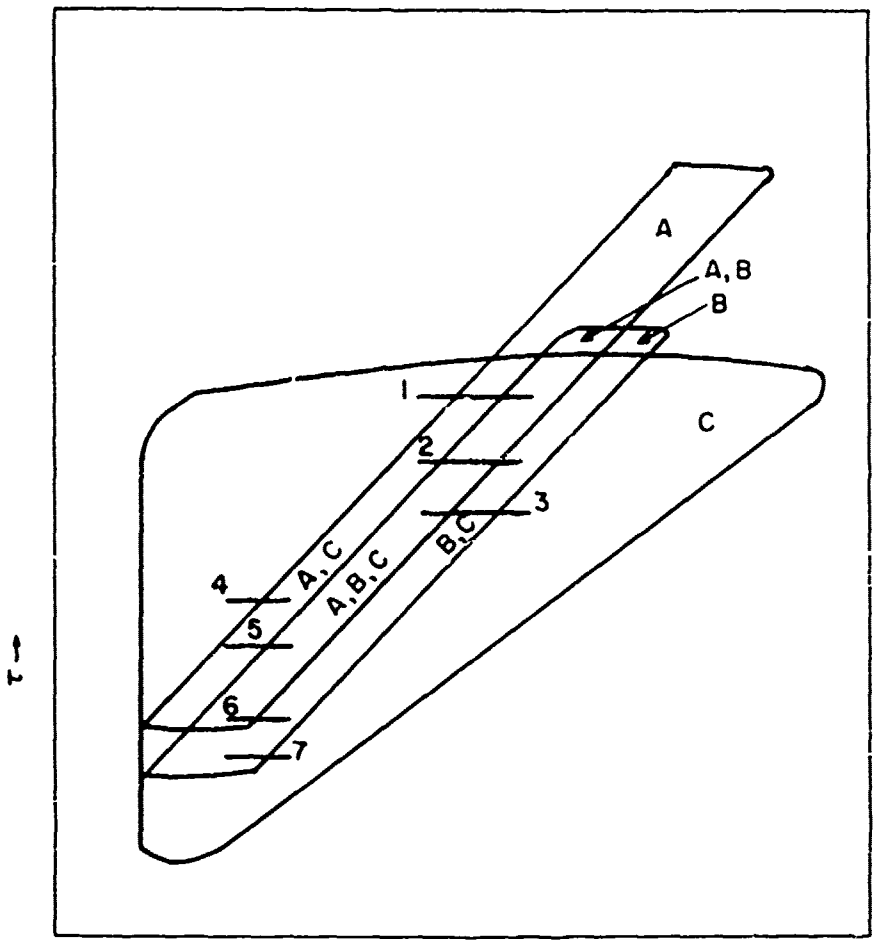


Figure 3-3. Ambiguities Caused by the Overlap of Three MORs. Traces 1,2,3,4 and 7 permit mode identification, while traces 5 and 6 do not.

The truncation of a trace at a boundary yields more definitive information because it uniquely associates the trace with the mode whose MOR is bounded. As an example, if Trace 5 were limited to the region ABC, instead of extending across the boundary into the region AC, we could conclude that the trace is the result of propagation via Mode B at whose MOR boundary the trace terminated.

In summary, the preceding discussion has identified three rules that can be applied to permit the indentification of propagation modes of radar detections that have been developed in the radar-observables space (τ-f plane). They are:

(1) Extension of the trace into (or total inclusion within) an unambiguous region indicates that the track resulted from propagation via the unambiguous regions mode.

(2) Termination of a trace at the boundary of an MOR indicates that the propagation occurred via the mode of the MOR in which the trace is contained. In applying this test, it is necessary to have made observations adequate to have disclosed an extension of the trace beyond the MOR's boundary, had it existed.

(3) Extension of a trace lying within a triply ambiguous region into two doubly ambiguous regions permits the identification of the propagation mode to be identified as that common to the two doubly ambiguous regions. (It is assumed that one and only one mode recurs in both of the doubly ambiguous regions.)

3.3 Probability of Detection and False Detection

The boundary tests for mode identification require that it be determined whether or not a target-generated trace extends beyond the boundary of the MOR in which it occurs. Because OTH radars are frequently operated at relatively low signal-to-noise ratios, there exists

some chance that any observation will give an erroneous result, indicating falsely the presence of a target or failing to indicate a target when an echo is indeed present. If the probability of such errors is greater than can be tolerated for the probability of misidentifying the mode responsible for a trace, it is necessary to defer the mode identification decision until a sufficient number of observations can be made to permit a more reliable decision. We here treat the probabilities of these decisions.

One approach to a tracking algorithm involves a predictive filter which, on the the basis of prior observations, extrapolates the observed target positions to a position at which it is anticipated that it will be detected in subsequent observations. As additional observations are made, the new information is used to improve the predictions for later observations. The logic for making the detections in such a tracker can take various forms. For example, a region of acceptance about the predicted point can be defined, and within that region the cell containing the largest energy can be regarded as the new detection. Because one cell will always have more energy than the others, this approach has the characteristic of making a detection each time and would be incapable of producing a "no detection" result. This property is undesirable for our present purpose because, as discussed above, the radar's failure to detect a tracked target in any given observation provides a certain kind of information required by the mode identification process.

For this reason it is important that the detection process employ some form of thresholding process. Cells with less energy than the threshold level are regarded as devoid of echoes. If two or more cells exceed the threshold, then the cell with the largest energy can be selected as the detection. A more elaborate scheme permits the threshold to be increased with increased distance of the cell from the point at which the echo's position has been predicted. In this way, detections are favored that most closely conform to the prior indications of the track.

The consequence of this kind of thresholding scheme is that there will be some probability of detection when echoes are present and some probability of false alarm when they are absent. There are many contingencies to the problem, aside from the commonly considered one of additive system noise with, say, a Gaussian distribution. For example, the predictive tracker may err because of past erroneous detection; or the target may accelerate, causing the echo to fall outside of the region of acceptance, or so far from the predicted position that its energy fails to exceed an increased threshold level. Also, polarization fading effects can cause the echo to be undetectable. For our present purpose, we assume that, in spite of these several contingencies, the detection process may be regarded simply as a Bernoulli process, i.e., a random process with two possible outcomes--detection or failure to detect--with probabilities P_D and $1 - P_D$, respectively. In the event that no echo is present in the cell, we may replace the probability of detection with the probability of false alarm so that the two outcomes of the Bernoulli process become P_{FA} and $1 - P_{FA}$. Then an M of N detection logic can be applied to determine the presence or absence of a track in any given subset of the data. In this way, a decision can be based on multiple observations, and the probabiltiy of error can be reduced correspondingly. However, the certainty with which decisions critical to mode idenfication can be made may be limited because they are based on a subset of all the observations of a given target--e.g., those in which detection in a region of reduced mode ambiguity is attempted. If the frequencies of the radar are selected so that there are too few attempts to observe echoes from the target in a critical region, a reliable decision is precluded.

It is desirable that the mode identification be made with a probability of correctness comparable to that required of the decision that a target is present and should be entered onto the display for the user of the radar. However, because the number of observations available in the region critical to mode identification is always a subset of the detections used for the decision to indicate the target on the display,

the probability of correct decision is always less. Because mode identification is necessary before the radar's coordinates can be converted to geographic coordinates, one must compromise by either

(1) tolerating a higher probability for error in mode identification than for target indication

(2) deferring display of the target until a reliable mode identification can be made

There are variations of these possibilities that can be used as well. For example, targets can be displayed at each possible one of the ranges when ambiguities exist or can be displayed prior to the time when a sufficiently certain mode identification can be made if the target appears to be within a geographical region where it poses a threat.

Let P_I be the probability with which it is required that the mode identification be correct and assume that the decision can be based on the detection, or failure to detect, a target in a subset of N observations. Also, let the probability of detection be P_D and the probability of false alarm be P_{FA}. Then, in the event that echoes are present in the critical region, the probability of realizing k detections in n attempts to detect is $b(k;n,P_D)$ where

$$b(k;n,P_D) = \binom{n}{k} P_D^{\,k} (1-P_D)^{n-k}$$

and the probability that there will be k_o or more detections is then

$$P[k \geq k_o;n,P_D] = \sum_{j=k}^{n} \binom{n}{j} P_D^{\,j} (1-P_D)^{n-j}$$

These are commonly used distributions, and can be found tabulated.* Similarly, the probability of experiencing k_o or more false alarms in n tries is

$$P[k \geq k_o;n,P_{FA}] = \sum_{j=k}^{n} \binom{n}{j} P_{FA}^{\,j} (1-P_{FA})^{n-j}$$

* for n ≤ 50, National Bureau of Standards, Tables of the Binomial Probability Distribution, Applied Mathematics Series, Vol. 6 (1950

Because an erroneous decision in the mode identification process has a similar effect, whether it be a false alarm or a missed detection, the "cost" of the two kinds of errors are equal--the mode operative in the detection is misidentified, and the target is interpreted as being in the wrong place; consequently, we wish to select values of k_o, P_D and P_{FA} such that the two kinds of errors have equal and suitably low probabilities. Assuming that the tests of regions containing echoes occur with the same frequency as tests made in regions where the echo is present, the probability of correctly identifying the mode P_I may be written:

$$P(k>k_o;n,P_D) = P_I$$

and $$P(k>k_o;n,P_{FA}) = 1 - P_I$$

In general, P_D and P_{FA} are parameters determined by the characteristics of the radar, and by the requirements placed on its probability of detection and false alarm. Particularly influencing these numbers would be the SNR of the radar echoes, the number of cells searched by the tracker, and the relative "cost" of false alarms and missed detections. Note that the probability of false alarm employed here, P_{FA}, applies to the overall process of attempting to extend the track with the current observation. Thus, if the tracking algorithm attempts to extend the track by seeking a detection (with uniform threshold) in any one of n cells, this probability is $n\,p_{FA}$ where p_{FA} is the more commonly encountered probability of a false detection resulting from a single trial. Typically, an OTH radar's tracker may search from 10 to 100 cells in its attempt to extend the track, so p_{FA} must be between one and two powers of 10 less than P_{FA}.

As an example, suppose that the signal-to-noise ratio at the detector is 10 db. Then, it should be possible to detect targets with a probability of 0.6 while experiencing a one-trial false alarm probability of about 10^{-4}. If the tracker examines 100 cells in its attempt to extend the track, the probability P_{FA} of erroneously extending a track when the signal is absent can be taken as 10^{-2}. In such a case, if as few as 2 detections in 10 trials are regarded as indicative of a target, the probability of successfully detecting

a trace extension is $P\{S \geq 2 \mid 10, .6\} \simeq 0.998$ while the probability of falsely detecting a path extension $P\{S \geq 2 \mid 10, .01\}$, has a value less than .0043, which means the probability of correctly concluding that no echo is present exceeds .995.

The above example demonstrates the power of a compound test of the m-out-of-n-trials kind. As expected, the test is less powerful if fewer trials are available for making the decision. For example, with the same probabilities of detection and false alarm, but with only five trials to make the decision, a criterion of 1 detection out of 5 trials would lead to a probability of correctly extending a trace of 0.913, and a probability of 0.951 that a correct decision not to extend the trace into an MOR would be made.

In the above examples, here only the incrementing of the number m of successes of a given n trials was at our disposal, the resulting probabilities of the two kinds of errors that can occur differed from one another somewhat. Nonetheless, the best choice that could be made in each of the cases resulted in error probabilities that would be quite satisfactory for many purposes. Figure 3-4 shows various combinations of P_{FA} and P_D which can be used with a corresponding m out of n decision criterion to achieve a probability of 0.95 that the existence of a trace in a critical MOR is correctly determined. Similar curves are readily developed for other values of P_I, using tables of binomial distributions.

3.4 Considerations for Testing at Soft Boundaries

At a "soft" boundary of an MOR, the signal-to-noise ratio gradually degrades until the signal becomes effectively non-detectable. The important difference is that there is no sharp boundary dividing regions where echoes arriving via a particular mode can be expected from those where they are unexpected. However, the methods described above can be applied if critical regions are correctly defined. Let us suppose a two-out-of-five decision criterion is being used to determine trace extensions into

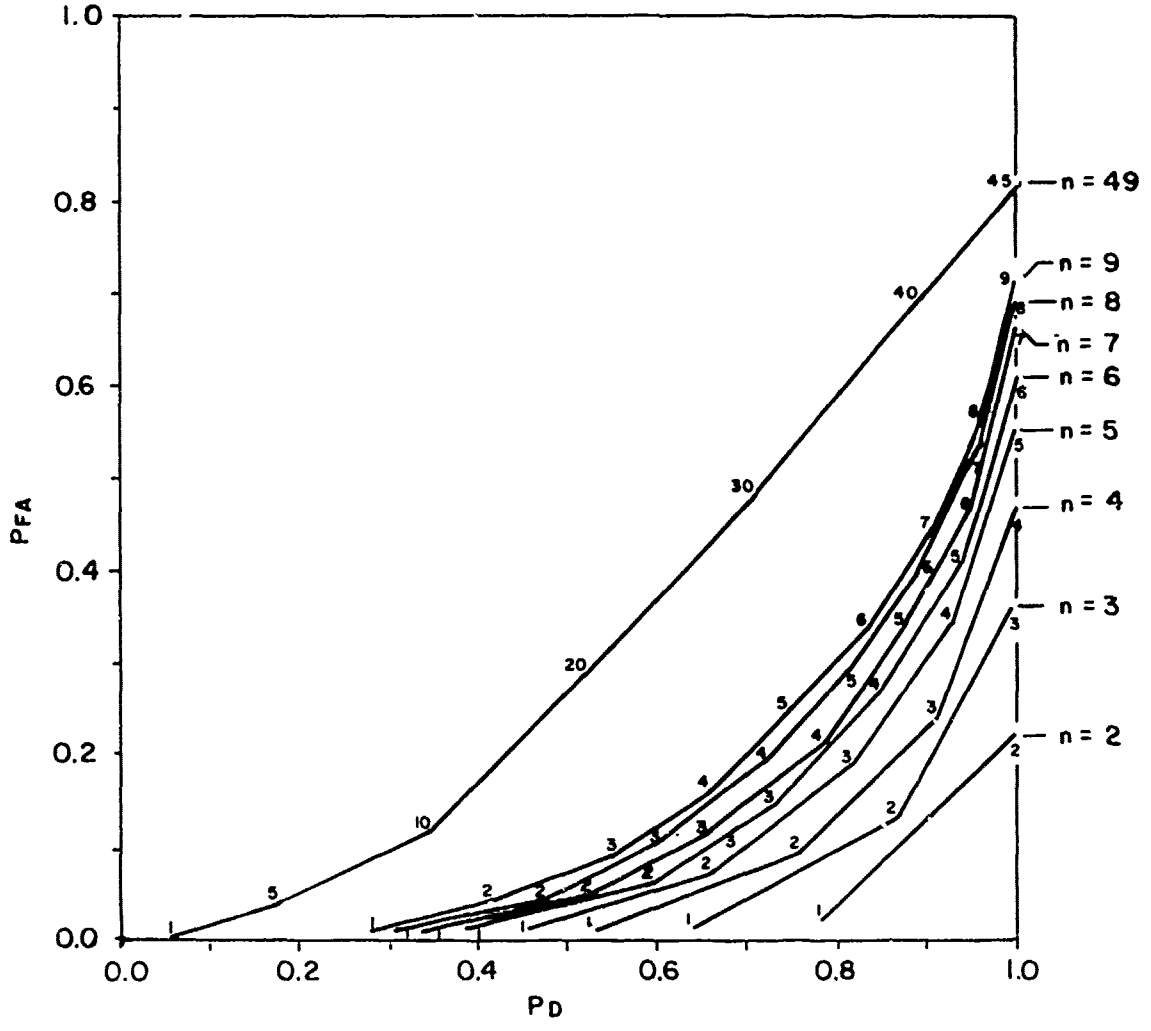


Figure 3-4. P_D and P_{FA} Required to Yield a Probability P_I=0.95 That the Existence of a Trace Extension into a Critical MOR is Correctly Determined for Various m out of n Criteria.

A75-182

critical regions and that the desired probability of correctly making these determinations is to be 0.95. Referring to Figure 3-4, we find that the probabilities P_D =0.66 and P_{FA} = 0.07 are required to achieve the required probability of a correct decision. A subset of five observations must be chosen such that we are sure of attaining this degree of certainty regarding the outcome of the individual observation. Let the softly bounded mode be called A, and let B be the mode with which the A mode is ambiguous. The critical region for the mode identification test must be chosen so that if the echoes have arrived via mode B, they will be detected by each observation with a probability P_D = 0.66. The signal-to-noise ratio of importance here is that of the B mode, which for now we assume to be uniform over the B-mode-observable region. The combination of B mode SNR and required detection probability determines an upper limit for the detection threshold T_D. In a simple detection problem a lower limit to the threshold value is fixed by the maximum false alarm rate which can be tolerated when no signal is present. A similar consideration applies here, except that the false alarm event is no longer caused solely by noise, but rather it can also be caused by a signal which differs from what would be a B-mode signal, only in having been attenuated by the mechanism that operates to bound the A-mode-observable region. In order to avoid a false mode identification using two of five decisions, we find by referring to Figure 3-4, that the probability of "false alarming" on a single observation must be .07 or less.

This single observation involves the testing of all points within some acceptance region determined by the characteristics of the tracker. Let us suppose that this acceptance region comprises N cells. (N will usually be somewhere between 10 and 100.) In N-1 of these cells, only noise is present, while in one of the cells there is the attenuated signal that has propagated via the A mode. The probability of making a detection based on noise only, we write as $P_{FA,N} \approx (N-1)\ p_{fa}$ where p_{fa} is the familiar probability of false detection for a single threshold comparison of

the signal sample from a single resolvable cell of the radar. The probability of the attenuated echo being detected in the remaining cell, denoted here P'_{fa} is just the probability of detection of a signal with the reduced SNR. Thus, the choice of the threshold for this decision is determined by the equation

$$P_{FA} = (N-1)\, p_{fa} + p'_{fa}$$

where, for the example under consideration, $P_{FA} = 0.07$. The relative magnitudes of the two terms will vary with the SNR of the attenuated A-mode signal. Figure 3-5 indicates how the required decision threshold must be varied with the anticipated value of the SNR for a given probability, p'_{fa}. Figure 3-6 indicates the minimum threshold levels that can be used to reject echoes of the specified SNR for the example. This chart was derived by (1) selecting a threshold level, (2) determining the probability of false alarm, p_{fa}, in 24 trials (assumed tracker search area), (3) subtracting p_{fa} from the desired overall probability of "false" detection, P_{FA}, to yield residual probability p'_{fa}, and (4) determining the SNR which would just lead to the residual probability p'_{fa} at the threshold level selected in (1).

The relationships among threshold levels, probability of detection, and SNR were determined by referring to Figure 3-5 and to graphs of the incomplete Toronto function published by Marcum (1960). The incomplete Toronto function $T_{\sqrt{v}}(1, 0, \sqrt{q})$ is equal to the probability of not detecting a single pulse signal of SNR = q using a threshold v (measured in units of twice the variance of the noise). Note that for our rather typical example, the threshold level required to reject a signal of SNR = 3 (4.8 db) with the desired probability is 4 standard deviations above the mean. At this threshold level (v = 8) in order to detect unattenuated (B-mode) targets with the desired probability of 0.66, it is necessary for their SNR to be 9.5 (9.7 db). In the absence of echoes, the probability of false alarming with a 4σ threshold would be 0.0003 on a single trial, or about 0.0075 in a 25-cell search area of the tracker. A threshold level of 3.5 standard

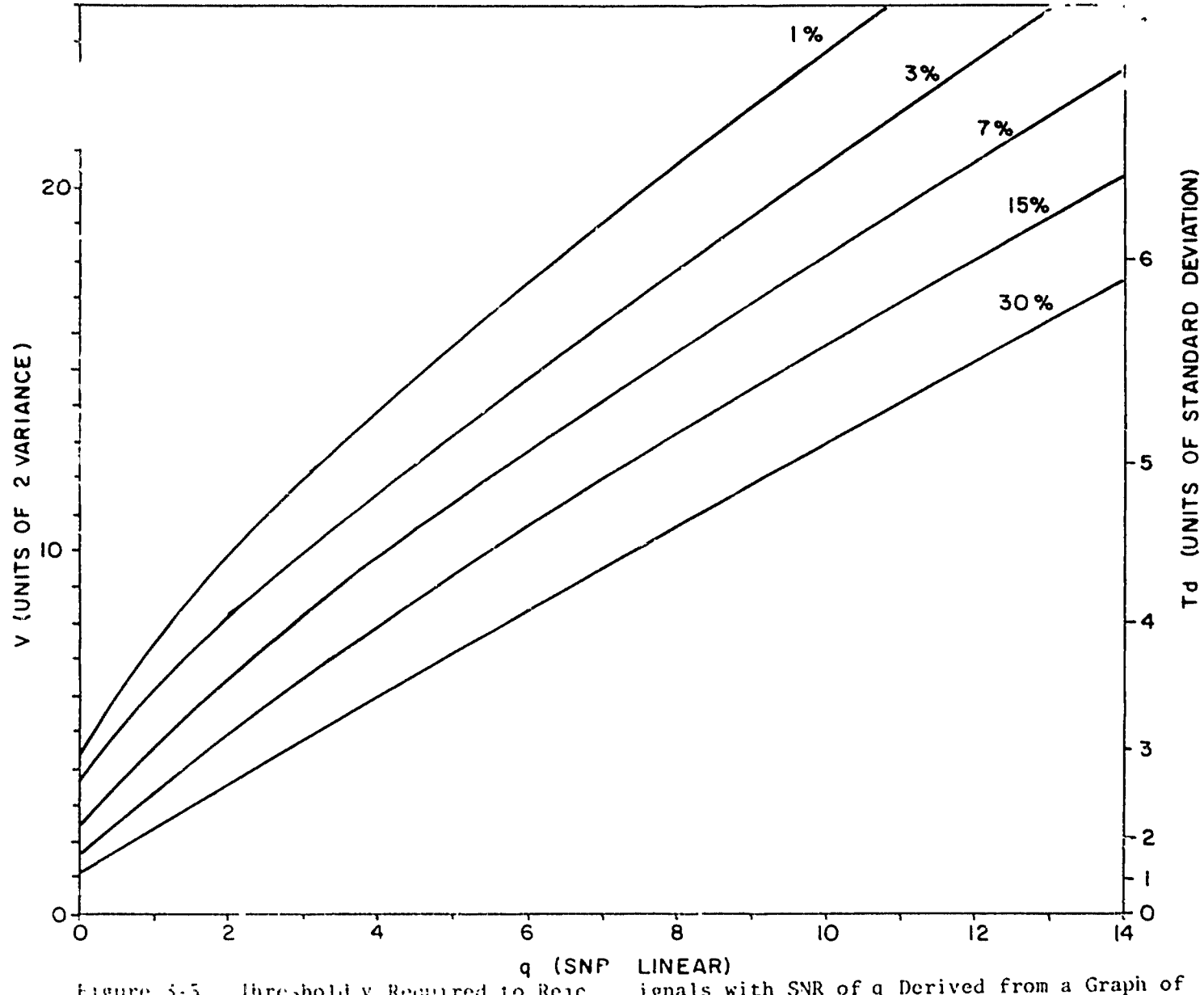


Figure 5-5. Threshold v Required to Reje ignals with SNR of q Derived from a Graph of the Incomplete Toronto Funct. $\int_{\sqrt{v}}^{\infty}(1, 0, \sqrt{q})$ (Marcum, 1960)

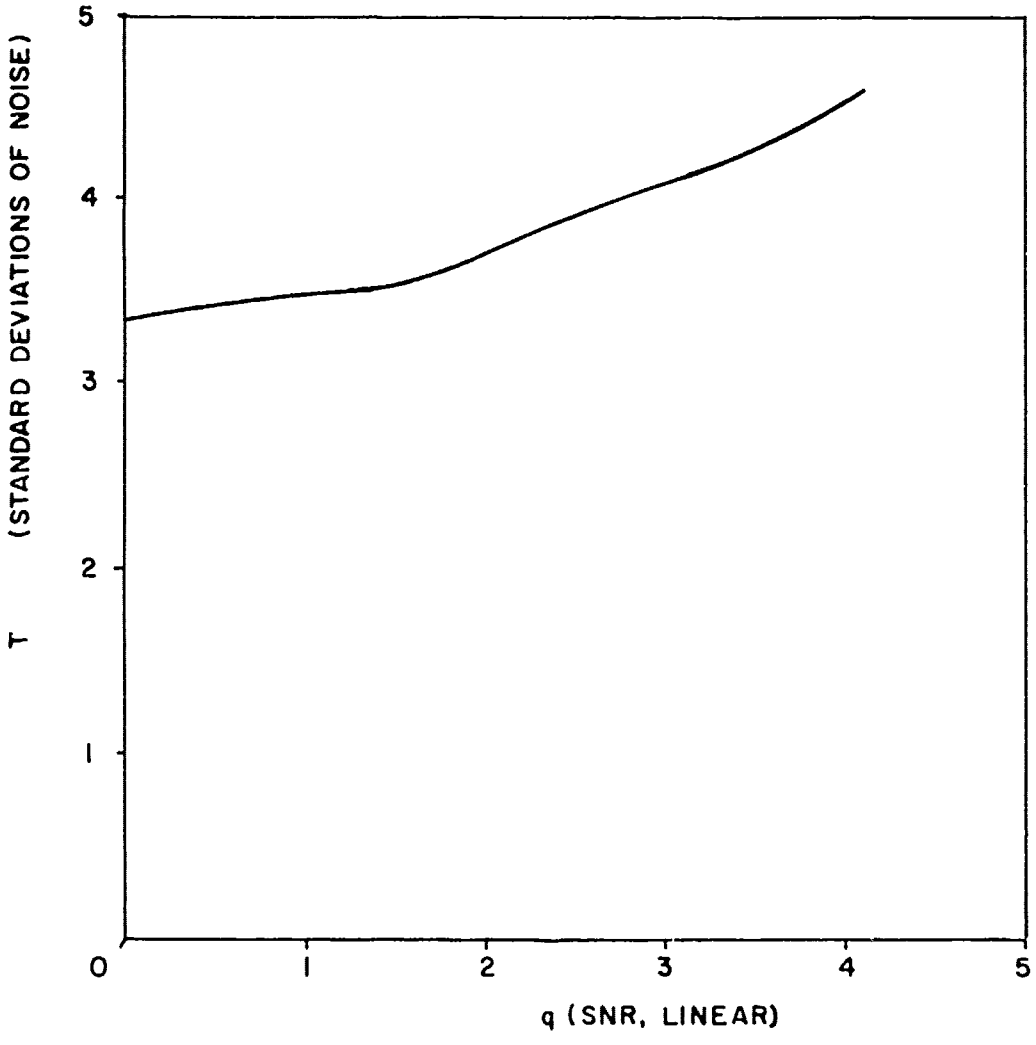


Figure 3-6. Maximum Threshold that can be Used to Assure that the Probability of Detection of the Attenuated Target, Plus the Probability of False Detection Within a 25 Cell Search Region is less than 0.07 (or 7%).

A75-184

deviations could be used if the signal-to-noise ratio of the A-mode (attenuated) echo could be assured to be no greater than unity (0 db). This second threshold level would permit a 0.66 probability of detecting echoes with an SNR of 7 (8.5 db) and would allow false alarms with a probability of .002 per trial, or .05 per search of a 25-cell search area. These probabilities are not far from those we initially assumed as reasonable for decisions involving a two-of-five decision criterion.

Summarizing the above discussion, we have considered the threshold levels required to determine whether or not echoes are present in the vicinity of a softly bounded MOR. For the example given, a two- out-of-five decision logic was used to obtain decisions with a 95% probability of correctness. A 4 σ threshold level was found to be required to reject echoes of SNR < 3 (4.8 db) with the required probability ($P_{FA} = 0.07$) and a threshold level of 3.5 σ was found necessary for SNR < 1 (<0 db). These threshold levels provide a 0.66 probability of detection (as required for the 95% probability of achieving two or more detections in five trials) for signals with SNR's of 9.7 db and 8.5 db, respectively. The probability of a false alarm caused by noise alone in a 25-cell tracker acceptance region was found to be 0.0075 and 0.05, respectively. Consequently, the use of the 4 σ threshold permitted the critical region of the B mode MOR to be defined as that lying outside the 4.8 db SNR contour of the A mode's MOR, but the use of this threshold level desensitizes the radar by about 1 db. If the 3.5 σ threshold is used, the critical region must be limited to the region outside the 0 db contour of A mode propagation, but there would result negligible desensitization of the radar.

The above example indicates how, by properly selecting the critical region of the f-τ plane for mode identification tests, satisfactory tests can be performed even at soft boundaries. The difficulty in the practical implementation of such tests is that of determining the SNR contours. Where it is necessary to determine the boundary with any degree of exactness, one must be pessimistic because there are numerous contingencies such as polarization, which will confound such predictions. It would be unwise

to attempt such mode identification tests, for example, to decide between high-ray and low-ray propagation unless the unambiguous region in which the test is to be made were extensive enough to permit testing well beyond the estimated position of the selected SNR contour of the high-ray MOR. In our earlier discussions of the extent of the MOR's, examples were given in which the extent of the F region MOR seemed to be insufficient for this purpose. This was particularly the case where an underlying layer obscured the F layer at frequencies only a little less than the MUF for any range.

3.5 Application of Tests to Fixed-Frequency Radar

The kinds of tests described in the above sections are easily applied to a frequency-hopping radar because the propagation to any given range via a particular mode is usually strongly frequency dependent. A radar that restricts its operation to a single frequency has only the change in range of moving targets to use as the variable in analyzing the mode of propagation. In general for each test made by varying the radar's operating frequency, there is a corresponding one that can be made by observing the echoes from a target which varies in range. Unfortunately, one of the strongest tests using frequency variation--that of testing at the MUF--becomes awkward when attempted with range variation. The passage of a target into a mode's skip region corresponds to increasing the frequency above the MUF to the target's range, and the strength of the conclusion that can be drawn is the same in both cases. However, the test has the unfortunate property of providing the information to determine the propagation mode at the very instant that the target may disappear from the radar's field of view. Because this is undesirable, it is preferable to attempt tests based on the first appearance of an echo at a boundary. Such tests are possible under the following propagation and target conditions.

(1) Receding targets at skip distance.

(2) Approaching targets at boundary of maximum range focusing (hard boundary).

(3) Approaching target at undetectable SNR contour (soft boundary).

In order to make valid tests of this kind, it is necessary that the radar's detection process be applied continuously to the critical region (in this case, to the critical radar delay intervals). Indeed, if an m of n decision criterion is to be applied, the testing must be carried out over a sufficient interval of radar delay in the critical region that the necessary n observations can be made of the moving targets. A similar radar delay interval of detection processing can be carried out just within the ambiguous region. Echoes first detected in the unambiguous region can then be tracked unambiguously into the region where ambiguity would otherwise exist. If, on the other hand, the echo is detected within the ambiguous region without first having been detected in the unambiguous region, it can be assigned to the second mode (that at whose MOR boundary the testing is performed). Figure 3-7 illustrates the procedure.

If the testing is to be performed at a soft boundary, considerations similar to those discussed in the prior section apply. The critical region must be limited to a region where the SNR of the attenuated mode is below the level used in the design of the decision process.

Clearly, this kind of mode identification procedure will suffer if track continuity cannot be maintained and so may only be feasible with trackers that operate with high reliability. This means the tracker must include adequate provision for carrying the track through the fades that commonly occur on HF signals.

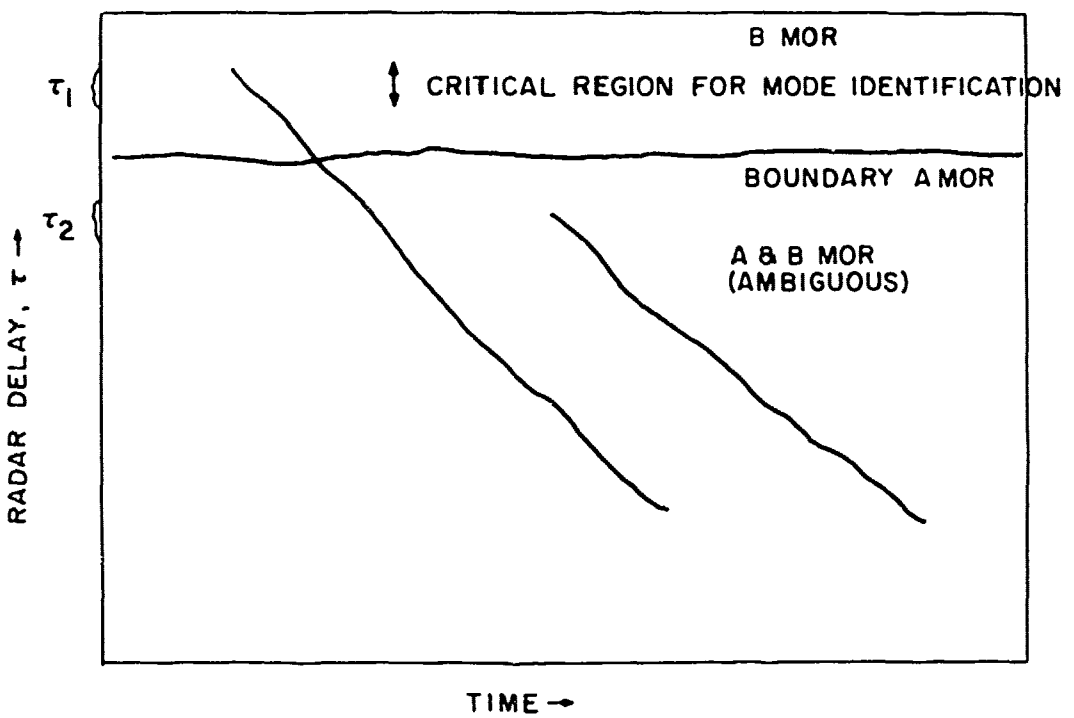


Figure 3-7. Use of Range Variation by Target in Mode Identification Test. Approaching targets first appear at delay interval τ_1 if B-mode supports propagation, and at delay interval τ_2 if A-mode supports propagation.

3.6 References

1. Marcum, J.I., "A Statistical Theory of Target Detection by Pulsed Radar: Mathematical Appendix," IRE Trans. Info. Theory, IT-6, 2, April 1960, Figure 14, p. 228.

4. AMPLITUDE FLUCTUATIONS AS AN INDICATOR OF PROPAGATION MODE

In the past, operational radars have used observed variations in the amplitude of the skywave backscatter as a means for sorting out propagation modes. The information gained by this process was used for propagation management, i.e., to insure that the region of the frequency-radar delay plane which was being displayed to the operator mapped to the geographical region of interest to the users. This kind of information could be extracted from the behavior of backscattered echoes from the earth's surface provided the characteristics of the ground clutter were at least, in part, determined by the propagation mode. For the purposes of this study, this kind of information has been assumed to be already available, and we have been concentrating on the question of how to identify the mode by which a particular target echo has propagated, when the echo appears in a region of the frequency-radar delay plane where it is known the echoes that have propagated via two or more modes can appear. It is the purpose of this section to examine whether each propagation mode imprints its own characteristic on the amplitude fluctuations observed on an echo from a discrete echo in the same manner as has been observed in the echoes from azimuthally and radially distributed ground clutter. If such a determination can be made, it would be of great utility for the identification of the mode operative in the observation of echoes with a fixed frequency radar, which otherwise only has the ability to test for mode identification when the target enters or leaves the geographical region covered by a given mode.

We first discuss the amplitude fluctuation effects which, in the past, have been observed and utilized for profile management. A discussion of the mechanisms responsible for the fading of ground clutter is next, followed by a discussion of how these mechanisms operate on a discrete target. Finally, a section discussing the observed fluctuations of signals from discrete scatterers is presented.

4.1 Use of Clutter Fading to Determine Mode Coverage

As a result of an extensive series of experiments performed in 1965-1966, it was found that the E and F modes could be distinguished by the fade rates that appeared on the ground clutter echoes. It was found that fixed-frequency observations indicated a fading period in the ground clutter that spanned a few seconds (1 to 5 seconds) when the F layer supported propagation and some tens of seconds (15 to 30 seconds) when the propagation was supported by the E layer. Thus, an examination of Z-axis records (echo strength displayed as recording density, as a function of radar delay and time) made with slow frequency scanning (dwells of approximately one minute on each frequency) permitted a determination of the regions of the frequency-radar delay plane in which either E-layer or F-layer propagation was dominant.

This method, when augmented with information extracted from an examination of meteor echoes and of discrete ground targets of opportunity provided the means by which radar profiling was accomplished without having recourse to a cooperative beacon. Although the methods have only been applied successfully to the problem of distinguishing between the modes supported by the sporadic E layer and the F layer, users of the method express optimism that it can similarly be made to work where the mode observable regions of the F1 and F2 layers are to be distinguished and in distinguishing among other combinations of modes as well.

4.2 Fading Mechanisms

Fading is a time variation in the amplitude of a received signal. Such time variations can be produced by two general classes of mechanisms. The first is a reduction of the signal power received because of attenuation of the signal introduced on the propagation path. The second general class of processes involves the interference among waves that have propagated over paths of differing phase lengths. A CW signal will undergo this interference fading because of

the combining of signals that have propagated via the differing propagation modes. However, here we presume such modes to have been time-, so this kind of multi-mode interference does not affect the amplitude of concern. However, there are several forms of multipath propagation which an OTH radar of moderate bandwidth (<0.5 MHz) does not resolve. Among these are the magneto-ionic multimoding and the so-called fine-grain multipath which results from ionospheric irregularities.

Also included here, under the category of interference effects, will be those caused by the interference of multiple, but unresolved, scatterers.

4.2.1 Fading Caused by Signal Attenuation (Energy Loss)

There are three mechanisms by which an ionospherically propagated signal can become significantly attenuated; these are absorption of the signal, defocussing of the signal, and scattering of the signal by interposed irregularities.

Absorption occurs predominantly in the D layer, through which HF skywave signals all must pass, regardless of their mode of propagation. This, along with the fact that the fluctuations in absorption occur rarely enough that their availability cannot be counted on for mode identifying tests, would seem to rule out their use. This is particularly the case for the sudden ionospheric disturbances (SID's) which are produced throughout the sunlit side of the earth by solar-flare-associated x-rays.

Ionization in the absorptive layers can also be produced by energetic particles, either in association with auroral phenomena or within the polar cap where low energy solar protons have access to the absorptive layer. Because of the spatial localization of the ionization so produced, the possibility exists that some kind of mode identification could be based upon complete knowledge of the spatial distribution of the absorption and of the ray paths of the propagation modes. However, we are a long way from possessing sufficiently detailed

information regarding the distribution of auroral ionization to make this approach feasible.

Defocusing of the signal can occur either because of a vertical distribution of an otherwise uniform electron density (as discussed in Section 2) or as the result of some horizontally structured perturbation of electron density such as can be produced by internal gravity waves in the atmosphere. The former mechanism would produce soft boundaries to mode-observable regions, as discussed earlier.

Internal gravity waves in the atmosphere can also produce electron density inhomogeneities with horizontal and vertical structure, which focus and defocus the radar's illumination. These effects will have characteristic periods of about fifteen minutes or longer. Irregularities associated with air motions of shorter periods have also been observed. Waves with periods less than five minutes are referred to as buoyancy waves or acoustic waves, the distinction being based on the relative magnitudes of various terms in the equation of motion. For our present purpose, we only note that these perturbations of electron density occur in the E and F layers and that the attendant focussing and defocussing is one of several causes of fluctuation to be discussed subsequently.

Signal strength can also be decreased by scattering. One phenomenon of this category is the obscuration by the sporadic E layer. Another is the similar obscuration caused by auroral ionization. However, when such an effect exists, a more noticeable signal fluctuation occurs through interference phenomena, to be discussed in a later section. Again, we note only that scattering is one of the phenomena that contribute to the fluctuation in signal strength.

4.2.2 Amplitude Fluctuation by Interference Effects

Interference effects are responsible for much of the signal strength fluctuation observed in the skywave signals. An OTH radar of moderate bandwidth (20 kHz) is immune to some of the effects encountered

in other services utilizing narrow bandwidths because the signals that have propagated via the gross modes (e.g., E layer, F1 layer, F layer high-ray, F layer low-ray) are time resolved and so do not interfere with one another. However, finer mode structure such as the magneto-ionic multimoding is often not resolved and so will manifest itself as an interference-caused fluctuation of signal strength. Irregularities are also a source of fine grain multipath, but their effect is more frequently analyzed in terms of a random diffracting screen (e.g., Booker, Ratcliffe, and Shinn, 1949). Extensive literature exists on this topic in which the fading is referred to as "scintillation."

The magneto-ionic multimoding occurs most strongly in signals that have propagated via the F-layer where the phase path through the magneto-ionic medium is very long. The irregularities have been observed at both E layer and F layer heights. The F layer irregularities would, of course, affect only F layer modes, while the E layer irregularities could introduce fluctuations on all modes.

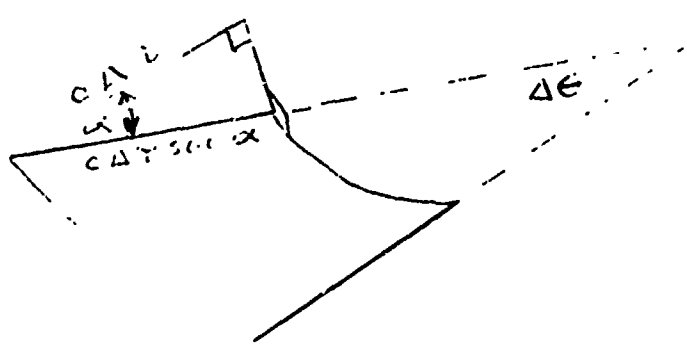
Time variations are required in each case. The magneto-ionic multimodes will change their relative phase as the layer shape, and consequently the ray path, geometry varies. Wind-induced drifts are usually invoked to explain the time variation of the irregularities in the E layer and magnetospheric electric fields to explain F-layer irregularity drifts.

These fluctuations may be regarded as having their origin entirely in the ionosphere. That is to say, they would occur regardless of the nature of the scattering target, be it a discrete or a distributed scatterer. We next take up some mechanisms for signal strength fluctuations which depend upon the nature of the target.

The first invokes time variations in the properties of the target to account for signal strength fluctuations. A target of sufficient size (several wavelengths) will have a scattering cross section that varies with the aspect from which it is viewed by a radar. Multiple, unresolved targets will scatter back echoes tha ay

interfere constructively or destructively, and so changes in their relative position will produce signal fluctuations. A similar effect occurs between a moving target and its image in the ground. Also, sea backscatter has a fluctuation as the approaching and receding Bragg-scattered components interfere, and as higher order scatterers interfere with each other and the first order components. The above mechanisms are all similar in requiring a motion of the scatterer to operate.

Consider next the backscatter from a distributed, but fixed, target such as from ground clutter. A pulsed radar with a time resolution leading to an interval of confusion $\Delta\tau$, and having azimuthal resolution $\Delta\theta$, responds to surface features by an amount determined by performing a vector integration over all scatterers in the illuminated field.



Thus, the area including the scatterers is

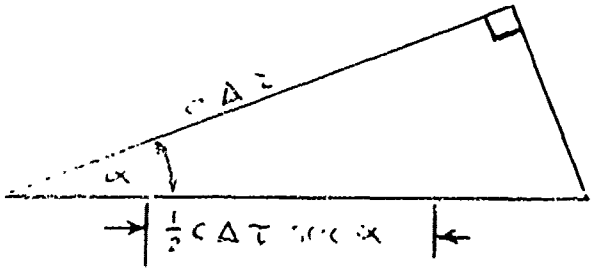
$$A\ (r,\ \phi) = r\ \Delta\theta\ c\ \Delta\tau\ \sec\alpha$$

N, the number of these scatterers is proportional to A. The summation over the set of scatterers in an unresolved area is mathematically equivalent to an N-step, two-dimensional random walk with the consequence

that the resulting amplitude has zero mean and a variance proportional to N. This variance is observed as the returned power. Thus, the use of a ground clutter coefficient δ_o with the illuminated area to determine the clutter's radar cross section is justified, and we can write the returned power or variance of the received amplitude as

$$P \propto \sigma_o c A r \Delta\theta \Delta\tau \sec \alpha$$

A further characterization of the clutter field is required for our purposes. This is an indication of the sensitivity of the clutter returned from any given range to changes in the angle of arrival of the radar illumination. Note that we are not interested here in the sec factor in the expression for the returned clutter power. Rather, we are interested in how quickly the random amplitude can change from a maximum value to a low value as the angle α is changed. This characteristic is measured by the autocovariance of the angular spectrum of the clutter returns from a resolved area. For our purpose we may obtain an estimate of this angular sensitivity by identifying the quantity with the resolution of an array with an aperture equal to the depth of the unresolved cell, $D = c\Delta\tau \sec\alpha$. The effective wavelength is half the electromagnetic wavelength because any displacement along the aperture affects the phase of the incident and backscattered waves.



The difference in phase between the signal of the front half of the cell to that of the back half is

$$\Delta\phi = 2\pi \frac{\frac{1}{2}c\ \Delta\tau\ \sec\alpha}{\frac{1}{2}\lambda}$$

$$= 2\pi\ f\ \Delta\tau\ \sec\alpha$$

Taking the derivative, we have

$$\frac{d\Delta\phi}{d\alpha} = 2\pi\ f\ \Delta t\ \sec\alpha\ \tan\alpha$$

And solving for the increment of α that will produce a phase change of π radians we obtain

$$\Delta\alpha = \left(\frac{d\Delta\phi}{d\alpha}\right)^{-1}\pi$$

$$= \frac{1}{2\ f\ \ \Delta\tau\ \sec\alpha\ \tan\alpha}$$

$$= \frac{\cos\alpha\ \cot\alpha}{2\ f\ \Delta\tau} = \frac{\cos^2\alpha}{2\ f\Delta\tau\sin\alpha}$$

Values of $\Delta\alpha$ are plotted for various frequencies and values of α in Table 4-1. A radar resolution, $\Delta\tau$, of 50 μs is assumed.

Table 4-1. Width of Angular Covariance (degrees)

f (MHz) / α	5	7	10	15	20	25	30
5	1.7	0.93	0.65	0.43	0.32	0.26	0.21
10	0.64	0.45	0.32	0.21	0.16	0.12	0.10
15	0.41	0.29	0.20	0.13	0.10	0.08	0.06
20	0.30	0.21	0.15	0.10	0.07	0.06	0.05
25	0.22	0.16	0.11	0.07	0.06	0.04	0.04
30	0.17	0.12	0.09	0.06	0.04	0.03	0.03
35	0.13	0.10	0.07	0.04	0 03	0.02	0.02
40	0.10	0.07	0.05	0.03	0.03	0.02	0.02
45	0.08	0.06	0.04	0.03	0.02	0.02	0.01
50	0.06	0.04	0.03	0.02	0.02	0.01	0.01

It is clear from Table 4-1 that rather small changes in the angle at which a resolvable cell is observed with the radar can produce large changes in the level of the radar echo. Also, note that the sensitivity to the angle of arrival decreases by more than an order of magnitude over the interval 5^{o} to 50^{o}. The implication of this is that for a given rate of ionospheric tilting, such an effect would lead to more rapid fluctuations of the ground clutter observed at near ranges than those observed at more distant ranges. Mitra (1949) presents some curves of the virtual height of the E, F1, and F2 layers, as functions of the time of day, from which can be scaled an estimate of the ionospheric curvature (change of layer tilt) in the east-west direction.

Let $h = a_o + b_1 x + c_1 x^2$

$$\frac{dh}{dx} = b_1 + 2c_1 x$$

$$\frac{d^2h}{dx^2} = 2c_1$$

where x is east-west distance and where a_o, b_1 and c_1 are appropriate expansion coefficients to fit the curve for Washington. D.C. summer presented by Mitra (1952).

Near sunrise the layer height descends at a rate of 50 km/hour. About two hours later, this rate has decreased to no more than 10 km/hour. Thus, we have

$$\frac{d^2h}{dx^2} \simeq \frac{40 \text{ km/hour}}{2 \text{ hours}} = \frac{20 \text{ km}}{\text{hour}^2}$$

At the latitude of Washington, an hour corresponds to a distance

$$x = \frac{1}{24} \cdot 2\pi \; r_e \cdot \cos L = 1348 \text{ km.}$$

where L is the latitude, and r_e is the radium of the earth. We may change the units of the above estimates of curvature to

$$\frac{d^2h}{dx^2} \left(\frac{\text{radians}}{\text{sec.}}\right) = \frac{20 \text{ km}}{\text{hour}^2} \cdot \frac{1 \text{ hour}}{1348 \text{ km}} \cdot \frac{\text{hour}}{3600 \text{ sec.}}$$

$$= 4 \times 10^{-6} \; \frac{\text{radians}}{\text{sec.}}$$

Or in units of degrees, about 2.30×10^{-4} degrees/second. Were this the only component of layer curvature, the characteristic time for a fade produced by this mechanism would be greater than 40 seconds.

$$\frac{0.01 \text{ degrees}}{2.3 \times 10^{-4} \text{ deg.sec.}^{-1}} \geq 40 \text{ seconds}$$

The gravity waves discussed earlier impose additional components of ionospheric structure that can lead to variations in the angle of arrival of the radar illumination on target. Clark, et al. (1971) have performed model calculations of the perturbation of the F layer electron density produced by gravity waves of various periods and wavelengths. As one example, the effects of an eastward propagating wave with a wavelength of about 300 km ($k = 0.02$ km^{-1}) and period of 20 minutes was computed. The amplitude of the wave was 2×10^{8} dyn$^{\frac{1}{2}}$ sec^{2}, which led to a wavelike deformation of the contours of equal electron density with peak-to-peak amplitudes of order 50 km. Idealizing the effect as that of producing a reflecting surface that deviates from a height of 300 km according to expression

$$h = A \cos k x$$

permits us to determine the shape of the layer, $\frac{dh}{dx}$. Thus,

$$\frac{dh}{dx} = - A k \sin k x$$

Then taking a time derivative produces the rate of change of slope.

$$\frac{d}{dt}\left(\frac{dh}{dx}\right) = - A k^2 \cos k x \frac{dx}{dt}$$

The gravity wave's velocity $\frac{dx}{dt}$ is equal to its length, Λ, divided by its period T. In this example, $\Lambda = 300$ km, and T = 20 min.

therefore

$$\frac{dx}{dt} = \frac{\Lambda}{T} = \frac{300 \text{ km}}{20 \text{ mins.}} = 15 \text{ km/minute}$$

The maximum magnitude of the rate of change of slope occurs when we have

$$\left|\frac{d}{dt}\left(\frac{dh}{dx}\right)\right|_{max} = A\, k^2 \frac{dx}{dt}$$

$$= 25 \text{ km} \times (0.02)^2 \text{ km}^{-2} \; 15 \frac{\text{km}}{\text{minute}}$$

$$= 0.15 \text{ radians/minute}$$

$$= 2.5 \times 10^{-3} \text{ radians/second}$$

$$\approx 0.143 \text{ degrees/second}$$

Referring to Table 4-1, we see that except at the lowest frequencies (≈5 MHz) incident at the lowest elevation angles (≈5°), an angular drift of this magnitude is sufficient to produce fading with periods less than a second. Clark, et al. (1971) computed the electron density perturbation produced by a model gravity wave with length about 1600 km and period 60 minutes. The amplitude of the gravity wave in this case was 2×10^9 dyn$^{\frac{1}{2}}$ sec^2. This wave produced a corrugation of the constant electron density contours of about 20 km peak to peak amplitude at the F layer. These values lead to a maximum rate of angular deviation

$$\left|\frac{d}{dt}\left(\frac{dh}{dx}\right)\right|_{max} = A\, k^2 \frac{dx}{dt}$$

$$= 10 \text{ km } (.004)^2 \text{ km}^{-2} \cdot \frac{1600}{60} \text{ km/minute}$$

$$= 0.00426 \text{ radians/minute}$$

$$= 0.004 \text{ degrees/second}$$

Again, referring to Table 4-1 we find that these waves could produce fading with a time scale measure in seconds or tens of seconds over much of the frequency and elevation angle interval of interest.

4.3 Fading Mechanisms as They Affect Echoes from Discrete Targets

The fading of echoes from a discrete target are of interest to us for two reasons. First, a tracker operating on an OTH radar must be able to bridge the periods of low detection probability that occur during fades. Secondly, we wish to determine whether the fading rate of the echo is characteristic of the propagation mode in the same manner as has been observed in the fading of ground clutter.

Not all the fading mechanisms discussed in the preceding section would contribute to the fading of a discrete target. Particularly to be excluded in this respect would be those that depend upon the target being distributed, i.e., those involving a change in the relative position of unresolved scatterers. Of those that do affect echoes from discrete targets, not all are capable of providing mode identifying information. Of this kind would be the fades caused by sudden ionospheric disturbances, which similarly affect all modes.

Some of the mechanisms are of a kind that would differently affect E-mode and F-mode signals. Magneto-ionic multimoding is one of these. For the E mode the change with range of the difference in phase between the ordinary and extraordinary modes is rather small, as manifested by the small Faraday rotation observed in E-mode signals (Earhart, 1973). That for the F mode, on the other hand, is large. Consequently, were this the sole source of fading in discrete target echoes, mode identification for major modes would be simple.

The effect of gravity waves is far more pronounced in the F layer than in the E layer; so again, we can anticipate a distinct difference according to propagation mode from this mechanism as well.

The remaining mechanisms are those associated with ionospheric irregularities. In most analyses, the effects of irregularities are regarded as random, as opposed to being analyzable into regular wavelike or periodic structures used to describe the magneto-ionic multimode interference (Faraday rotation) and the focusing effects of traveling ionospheric

disturbances. Also, the structure of the effects is rather different. While irregularities structure the radar's illumination fields into interference-caused "speckles," measuring a kilometer or less, the other mechanisms cause focusing or polarization bands that measure tens or hundreds of kilometers in width. If the effects are superposed, the irregularity fading is apt to mask the other effects.

The theory of ionospheric irregularities that cause signal scintillation is not as well developed as those for Faraday rotation and traveling ionospheric disturbances. Particularly lacking are theories which can predict differences in the scale size or probability of occurrence that may exist between E-layer and F-layer irregularities. For this reason, the following section examines empirically the difference in fading rates that occur in E-and F-layer modes.

4.4 Empirical Evidence Regarding Usefulness of Target Fading in Mode Identification

Although abundant literature exists on the topic of fading HF signals, the present study must be restricted to those effects which persist when the time resolution of the probing signal is sufficient to resolve the multiple propagation modes. That is to say, we must exclude from consideration those effects which are themselves manifestations of multimode propagation that would be resolved by the radar.

Experimental evidence has been collected by sounding on closely spaced frequencies, indicating that signal fading is associated with effects that occur near the reflection height. (Essex and Hibberd, 1968, and Essex, 1968). Thus, if the physical processes responsible for the observed echo fading were significantly different for the different regions of the ionosphere, it would be reasonable to expect that each region imparts a characteristic signature or modulation on a signal which could be used for mode identification.

However, the use of the fading characteristic of the target for mode identification is complicated by the fact that the fading is usually the result of time variations of a spatially structured illumination pattern. Thus, motion of the target would modify the effects that have been observed over propagation paths with fixed end points. If the target were motionless or moved sufficiently slowly, the characteristics observed by fixed end-point sounding would be immediately applicable. On the other hand, if the targets moved sufficiently rapidly, the expected fading characteristics could be predicted by considering the rate at which the target passes through an essentially stationary or "frozen" illumination pattern. Considerably greater difficulty is to be expected when the fade rates associated with changes in the propagation path are comparable to those caused by the motion of the target through the spatially structured illumination. As we shall see, this latter case is all too often encountered.

In this section the means and requirements for performing mode identification are discussed. The basis of this discussion is this Laboratory's experience and the information available from published literature. Many mechanisms have been proposed to describe and characterize the observed fading effects. Some of these mechanisms have been discussed previously. The purpose of this section is not to validate the theories and mechanisms invoked to explain the observations of signal fading; instead the purpose is to discuss and indicate possible ways in which the fading of the signal might be applied to solve the problem of mode identification.

4.4.1 Background and Measurement Techniques Used to Characterize Amplitude Fading of Signals

The temporal variation of signals can be determined by measuring the received amplitude of each pulse transmitted over the path. The temporal variation or fading of ionospherically propagated signals has been investigated extensively, dating back to Ratcliffe and Pawsey (1933). Considered here are only those observations made with pulse length short enough (less than 200 usec.) that the transmissions via the different modes are resolved. The ordinary and extraordinary components can be separated on the basis of time-of-arrival or delay time with sounders having time resolutions of one microsecond order. Additionally, a number of investigators utilized circularly polarized antennas to further ensure that the cause of the fading was not Faraday rotation.

Although fading produced by Faraday rotation could be used as an indicator of propagation mode, this section stresses fading caused by small scale ionospheric irregularities which are changing with time and might be different for the different modes. These effects occur frequently enough and obscure the more predictable Faraday rotation so that they must be contended with in the development of a mode identification procedure. As indicated, some of the other causes of the signal fading can be eliminated by proper choice of operating parameters so that the cause of fading can be limited largely to those cases of interest.

Fading effects have been characterized and classified in various ways. Of particular interest to us are:

- o Fading frequency spectrum or fade rate.
- o Amplitude probability density function

Each of these methods has advantages and offers a different dimension for studying the amplitude fading of a signal. The spectrum analysis separates the fading rate into the various fading frequency components. Sometimes a crude estimate of the fading spectrum is made by counting the number of observed amplitude maxima (or minima) per unit of time.

The amplitude probability density function is typically studied to identify which type of distribution function, such as Rayleigh, Rice, or Gaussian, characterizes a set of measured amplitudes. The form of the distribution function, besides being of interest in communications system design, implies certain physical characteristics about the properties of the transmission medium such as the strength and multiplicity of the scattered components.

4.1.2 Previous Work

The reported data on signal fading exists for two different sounding techniques, one using essentially vertical incidence propagation, and the other oblique incidence propagation. Evidence (McNicol, R., 1949, and Kokourov, et al., 1971) indicates that the effects measured on oblique paths are different from those measured simultaneously on a vertical path. Of interest for the current application to an OTH radar is, of course, an oblique path. Unfortunately, the larger quantity of data available is for the vertical incidence propagation case.

The large volume of vertical incidence data was collected to measure the motion of ionospheric winds and the structure of the wind-driven moving irregularities. These investigations usually follow the simple method of spaced receiver by Mitra (1949) or the more

sophisticated techniques applied by Booker, Ratcliffe, and Shinn (1950) or Briggs, Phillips, and Shinn (1950). Although not directly applicable, it is possible to gain insight from the extensive amount of data collected using vertical soundings.

The case of vertical incidence propagation is discussed first in the following section. The next section discusses the data available on fading for the oblique propagation case.

4.4.2.1 Fading Characteristics of Signals Observed at Vertical Incidence

Amplitude distributions were measured as a function of mode on frequencies of 2 and 4 MHz by Flood (1954). The amplitude fading was categorized for each observation as being either a Rayleigh or a Gaussian distribution. The data were then tabulated and the percentage of total observations having each type of distribution was reported as a function of propagation mode as shown below:

MODE	PERCENTAGE OF OBSERVATIONS WITH	
	RAYLEIGH DISTRIBUTION	GAUSSIAN DISTRIBUTION
E	59%	41%
F	48%	52%

Although there is some difference in the predominant type of distribution for the two modes, it can be seen from this data that the differences are too small to provide a means of mode identification.

Data on the fade rates is often not reported directly, although such data have been used to deduce the results indicated. The fade rate can be constructed from quantities reported in the literature by combining the reported scale size* and drift velocities of the irregularities.

A large volume of data has been reported on the measured drift velocity of the ionospheric irregularities. The measured velocity of the winds varies with a number of geophysical factors including season

*an estimate of mean size

and time of day. The reported wind speed measurements generally fall in the range of 40 to 150 m/sec. Davies (1969) quotes the average component of the wind speed as being 80 m/sec in the E region and about 100 m/sec in the F region. Rao and Rao (1957) measured a wind speed of 84m/sec in the E region in agreement with the reported results of Davies. Mitra (1949) reports the most frequent wind velocity observed was 50 m/sec. Mitra does not segregate the data by ionospheric propagation mode. However, he specifically conducted a series of tests designed to compare the wind speed in the E region with that in the F region. This was done by collecting alternate samples from each region using an alternating delay time gate. On the basis of a large number of these samples, Mitra concluded that the magnitude and also the direction of "wind was always nearly the same for the two echoes" from the different regions.

A more recent measurement of the ionospheric drifts was reported by Kokourov et al. (1971); some of their histograms are shown in Figure 4-1. The data are shown for upper ionosphere (F region) and lower ionosphere as measured at two locations in the Soviet Union as a function of season. The most frequently observed drift velocities, between 40 to 120 m/sec. are essentially in agreement with previous reported data. The important point to be made is that again there is no large difference in the observed drift speed between the E and F regions of the ionopshere. This implies that the fade rate will be different in the two regions only if the scale sizes of the irregularities are significantly different.

Histograms of the frequency of occurence of different values of the scale size of the irregularities are shown in Figure 4-2 which is taken from Essex and Hibberd (1968). This data indicates that the mean scale size of the E region and F region irregularities differ by a factor of 2 to 1. The measured scale size of the sporadic E irregularities is between the values measured for the F region and E region. Also, some additional data were collected on the Buckland Park array in Australia

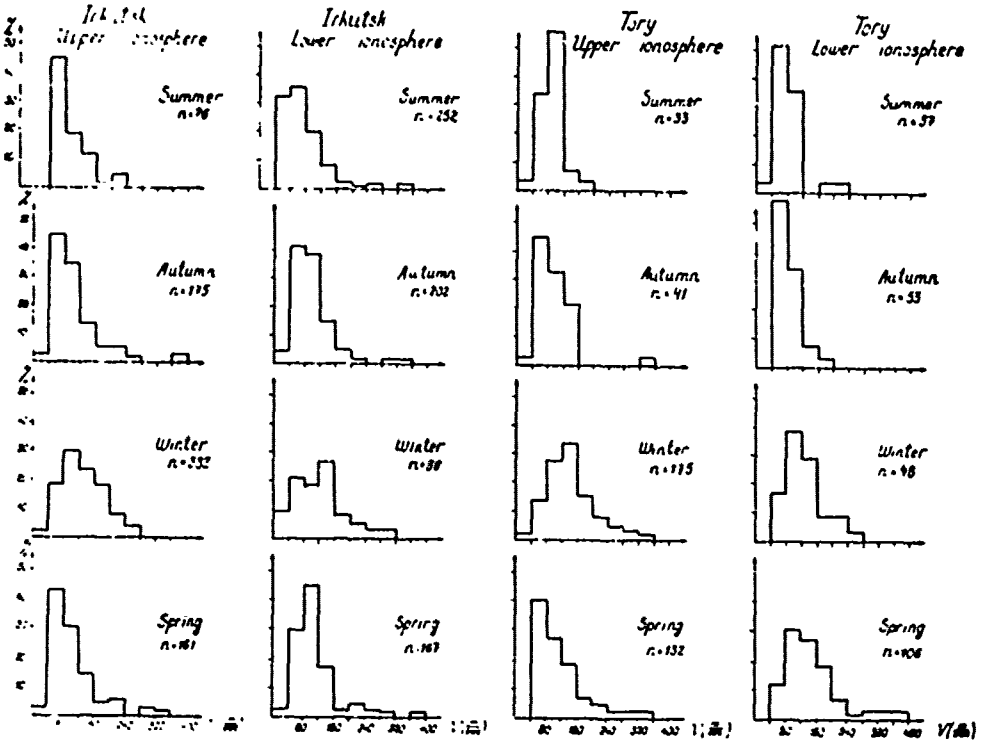


Figure 4-1. Distributions of Drift Speed

Reproduced from Kokourov, et al. (1971)

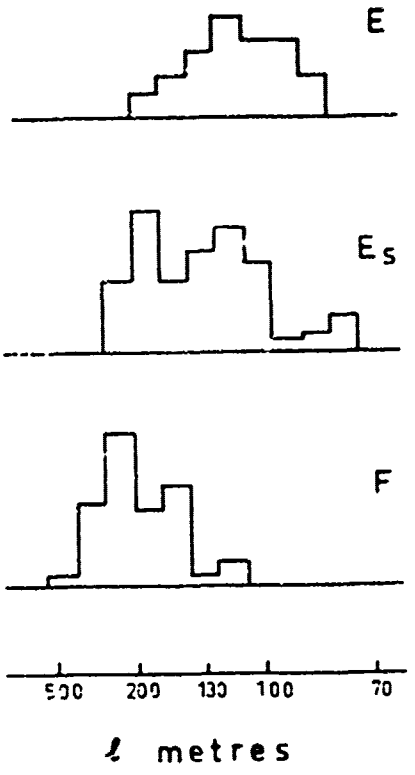


Figure 4-2. Histograms of the Frequency of Occurrence of Different Values of the Horizontal Scale Size, ℓ, of the Ionospheric Irregularities

Reproduced from Essex and Hibberd (1968)

by Felgate and Golley (1971). The array consists of 80 independent aerials. The output display consists of light pattern formed by arraying a series of lights in the same manner as are the antennas and driving the output of each light in proportion to the signal strength observed in the corresponding element of the array. The observed light patterns were grouped into seven classes that included large-scale, fringes, crossing fringes, random, and no fading. The percentage of time that each pattern was observed was reported. A comparison of these results for the large-scale patterns and the case of no fading is shown for each mode in the table below.

PATTERN TYPES	PERCENTAGE OF OBSERVATIONS FOR THE INDICATED PATTERN TYPES		
	E	E_s	F
Large Scale	12	19	46
No Fading	2	4	24

The table indicates that the fading of the signal reflected from the F layer was categorized as large-scale or non-existent 70% of the time while that from the E and E_s layers was so categorized only 10 to 20% of the time. By large-scale the authors meant fading structure that had a scale larger than their 1 km antenna array diameter. Thus, the spatial structure of the illumination was observed to have a scale less than one kilometer in size 30% of the time when the F layer produced the reflection and in excess of 80% of the time when the E layer produced the reflection. However, the size scales for the two modes do overlap, and inasmuch as the mode identification requires a difference discernible in a single sample (i.e., the fading on a single track) rather than being discernible in a statistical analysis, these results are not conclusively favorable.

4.4.2.2 Fading Characteristics of Signals Observed at Oblique Incidence

As previously indicated, there is less data available on fading statistics for the oblique propagation case. There are, however, some limited results available from the POLAR CAP III program and the POLAR FOX I program that, although not extensive, are valuable because they are directly applicable to the problem in hand.

In the POLAR CAP III program, a series of relative power measurements were made of echoes from a remote beacon located about 1600 km from the radar. The final report (ITT-EPL Staff, 1974). discusses the results of two case studies: one when the propagation was by a sporadic E layer and the other when the propagation was by an F layer. The results are shown for a 32-pulse dwell (one second) for both cases in Figure 4-3 and 4-4. The outputs shown are fully processed data which included correlation and Fourier transformation for doppler detection. Each output is an independent sample taken at the doppler offset frequency of the target as observed on an amplitude versus doppler frequency display. The sampled output was collected at a point in the processing chain at which inputs for a mode identification algorithm would be supplied.

A total of 16 seconds of data are shown in Figures 3 and 4. The important feature to note is the large variation (18 db) of the relative amplitudes for the F mode case in comparison to that (6 db) for the E mode case. It should be emphasized that these data were collected in the polar environment and represents a very limited quantity of data. There is no means to interpret statistically how applicable this data would be to the non-polar region. This data merely illustrates the type of amplitude variation differences necessary to perform mode identification.

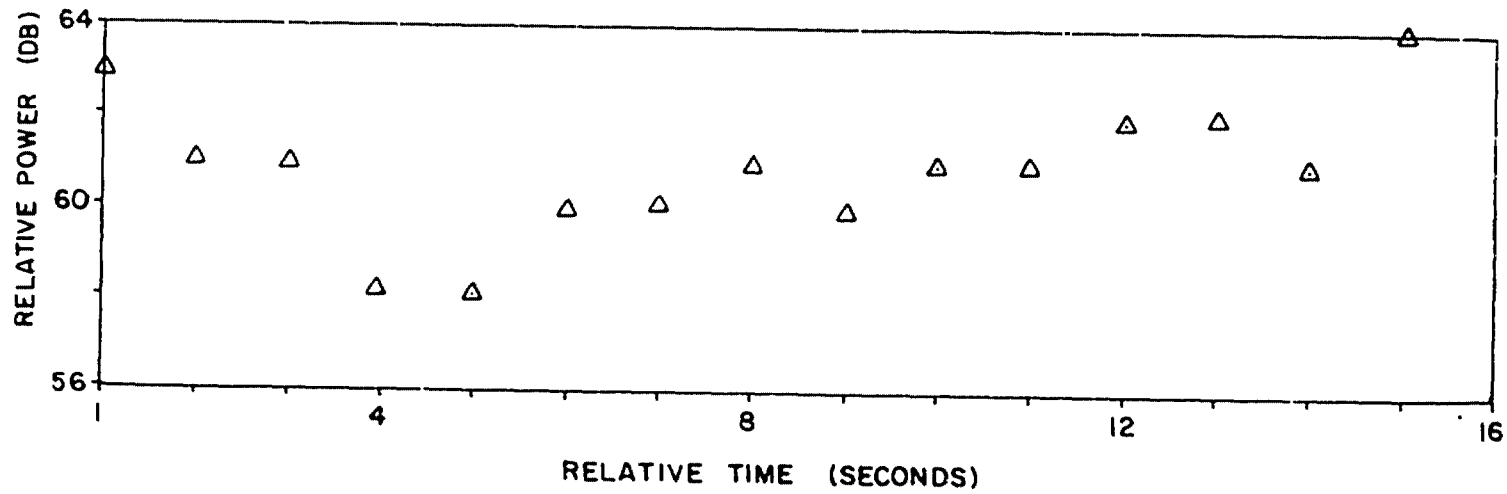


Figure 4-3. Relative Received Power of the Upper Sideband for Each 32-Pulse Dwell (1 Second), 6 July 1973, E Mode Propagation

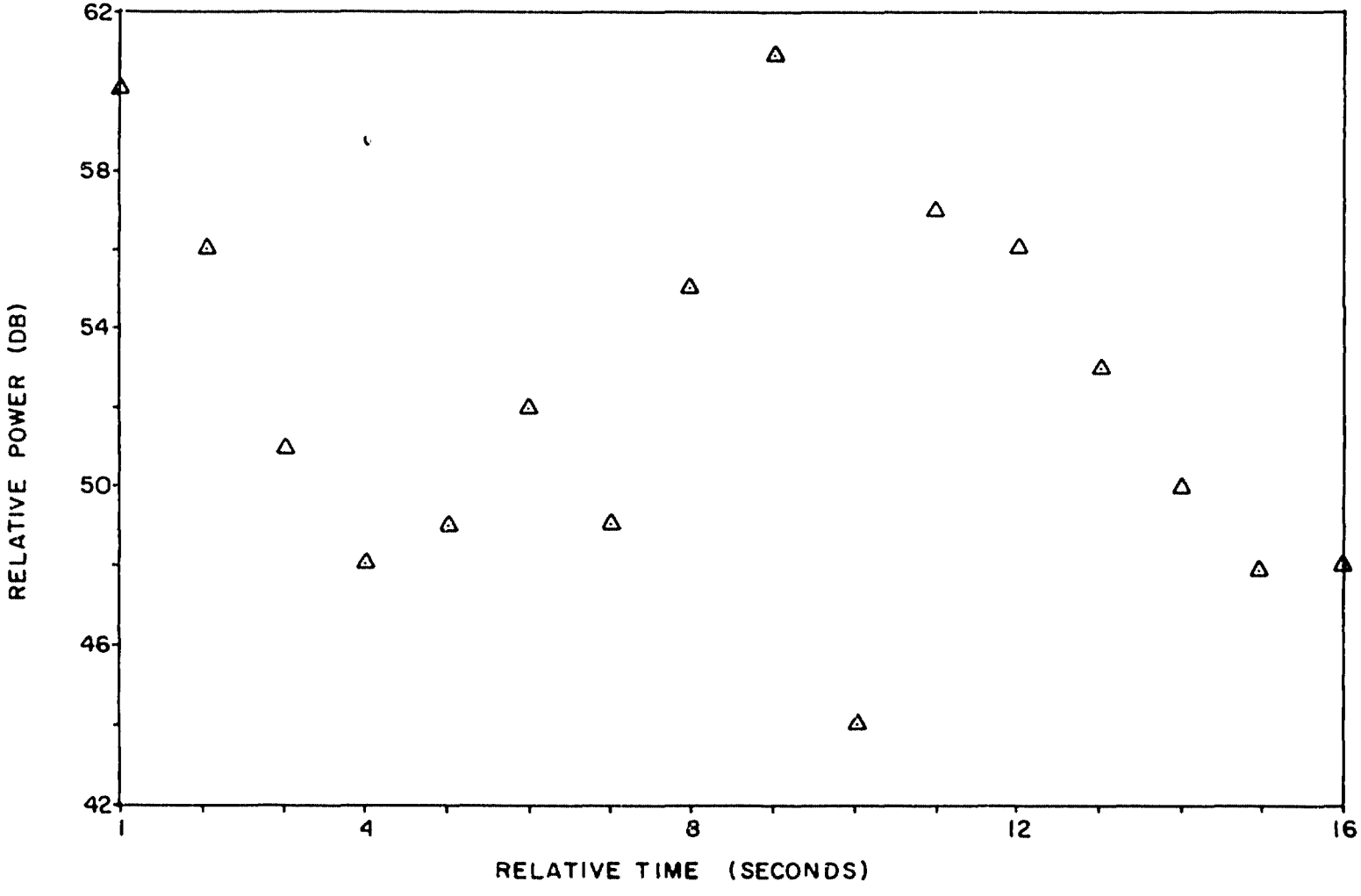


Figure 4-4 Relative Received Power of the Upper Sideband for Each 32-Pulse Dwell (1 Second) 1 December 1973, F Mode Propagation

In the data discussed up to this point, the terminals used for reception and transmission have been located at fixed position. Also, in the preceding examples the beacon was at fixed land position. A typical airspeed, say for a Boeing 707 aircraft, is around 400 mph (650 kmh) or 177 msec. The speed of the aircraft is only slightly larger than the previously indicated measurements of the wind speed within the ionopshere. Thus, the fading characteristics of echoes from an aircraft target could be influenced significantly by the motion of the aircraft.

A comparison of the fading statistics of aircraft and repeater targets was made by Riverside Research Institute (RRI) and was reported at the 1972 OHD Technical Review Meeting (Mastrandrea and Healy,1972). These data were collected with a Whitehouse radar during the POLAR FOX I experiments. The reported data show amplitude as a function of time and cumulative distributions of the amplitudes. The beacon was located at a nominal range of 1930 km from the radar. Two aircraft targets were observed along the heading of the beacon for the cases shown. The observed shape of the distribution was only slightly different (largest reported difference was 20%) for the combined results of the two aircraft and the beacon targets. This would appear to be relatively insignificant in view of the limited data base. (The authors emphasized that " . . . . . the sizes of the data samples involved are small.") The variations of the distribution for the two aircraft targets appears to be on the order of 20%.

The amplitude vs. time histories were presented for the repeater and aircraft over a total duration of 14 minutes. During portions of this period the repeater and aircraft target were present simultaneously. A visual inspection of this data did not indicate any significant differences in the characteristic of the amplitude variation of the repeater and the aircraft echoes. In fact, during the major portion of the time when the echoes were observed simultaneously,(which extended over a period of some 3 minutes) the two sets of echoes appeared to fade with similar periods.

On the basis of these results, it would appear that the motion of the aircraft had little effect on the fading characteristics observed by an OTH radar. This would imply either that the observed fading represents a temporal variation with relatively little spatial structure or that the fading rate was too rapid to be unambiguously represented by the one-second samples taken with the radar--so that the observed fade rate is system, rather than phenomenon, limited.

4.5 Intermediate Summary and Conclusions

To summarize the investigation of the possibility of using fading characteristics as a propagation mode discriminant, the following points may be made.

1. The amplitude distributions that have been observed are similar for signals propagated via the E and F modes.

2. Irregularity drift measurements indicate similar drift (or wind) speeds occur in the two layers.

3. The mean irregularity size deduced from spaced receiver measurements indicate that the F layer irregularities have a mean scale size about twice that of the E layer but that there is a substantial overlap in the size distributions that were observed.

4. Irregularity-induced fading occurs a substantial fraction of the time (>50% of the time in the Buckland Park data) and a practicable mode identification technique must be operable in their presence; this precludes complete dependence upon the Faraday rotation effect for mode identification.

We may conclude from the above that although the fading produced by ionospheric irregularities on signals may differ somewhat in scale, size and frequency of occurrence, depending upon the ionospheric layer supporting the propagation of the signal, the differences are statistical rather than deterministic. For this reason, the fluctuation of signal amplitudes do not, of themselves, provide a conclusive indication of the mode supporting propagation of echoes from a given target. Thus, while amplitude fluctuations provide a diagnostic means that may be helpful in determining mode structure (i.e., the distribution of mode-observable regions in the f-τ plane) they are of little value by themselves in determining the propagation mode of specific detections.

4.6 References

1. Booker, H.G., J.A. Ratcliffe and D.H. Shinn, "Diffraction from an Irregular Screen with Applications to Ionospheric Problems," Phil. Trans. Roy. Soc., A, 242, London, 1950, p. 106.

2. Briggs, B.H., G.J. Phillips and D.H. Shinn, "The Analysis of Observations on Spaced Receivers of the Fading of Radio Signals," Proc. Phys. Cos., B, 63, London, 1950, pp. 106-121.

3. Clark, R.M., K.C. Yeh and C.H. Liu, "Interaction of Internal Gravity Waves with the Ionospheric F2-Layer," J. Atmos. Terr. Phys., 33, 10, October 1971, pp. 1567-1576.

4. Davies, K., Ionospheric Radio Waves, Blaisdell Publishing Corp., Waltham, Massachusetts, 1969, p. 370.

5. Earhart, J.R., "The Ordinary-Extraordinary Mode Clutter Cancellation Study (U)," CHAPEL BELL Report No. 110, ITT Electro-Physics Laboratories, Inc., Columbia, Maryland, June 1973 (SECRET).

6. Essex, E.A., "Periodic Fading of Ionospheric Echoes," J. Atmos. Terr. Phys., 30, 7, July 1968, pp. 1441-1443.

7. Essex, E.A., and F.H. Hibberd, "Frequency and Spatial Correlations of Fading Radio Exhoes from the Ionosphere," J. Atmos. Terr. Phys., 30, 7, May 1968, pp. 1019-1031.

8. Felgate, D.G., and M.G. Golley, "Ionospheric Irregularities and Movements Observed with a Large Aerial Array," J. Atmos. Terr. Phys., 33, 9, September 1971, pp. 1353-1369.

9. Flood, W.A., "The Fading of Ionospheric Signals," Technical Report No. 17, Cornell University, August 1954.

10. ITT-EPL Staff, "POLAR CAP III Final Experiment Report," Project Report No. 262, ITT Electro-Physics Laboratories, Inc., Columbia, Maryland, September 1974 (SECRET).

11. Kokourov, V.D., E.S. Kazimirovsky, V.N. Zakharov and E.I. Jovty, "The Simultaneous Measurement of Ionospheric Drifts by Vertical and Oblique Sounding," J. Atmos. Terr. Phys., 33, 6, June 1971, pp. 943-950.

12. Mastrandrea, R., and C. Healy, "A Comparison of the Fading Statistics of Aircraft and Repeater Targets (U)," OHD Technical Review Meeting of 3-4 May 1972(a), Vol. III, May 1972 (SECRET).

13. McNicol, R., "The Fading of Radio Waves of Medium and High Frequencies," Institute of Electrical Engineering Procedures, Part 3, Vol. 96, October 1949, pp. 517-524.

14. Mitra, S.N., "A Radio Method of Measuring Winds in the Ionosphere," Proc. I.E.E.E., III, 96, 1949, pp. 441-446

15. Mitra, S.N., The Upper Atmosphere, The Asiatic Society, Calcutta, India, 1952, p. 249.

16. Rao, M.S., and B.R. Rao, "Analysis of Fading Records from Four Spaced Receivers for Ionospheric Wind Measurements," J. Atmos. Terr. Phys., 10, 1957, pp. 307-317.

17. Ratcliffe, J.H., and P.L. Pawsey, Proc. Camb. Phil. Soc., 29, 1933 p. 301.

5. ALGORITHM FOR MODE-IDENTIFYING TRACKERS

This section discusses the incorporation of the previously discussed mode identification criteria into an automatic tracking algorithm. These considerations include the provision of necessary information regarding mode structure to the tracker, the information to be retained by the tracker concerning mode structure, and the logic of the decision processes to be applied.

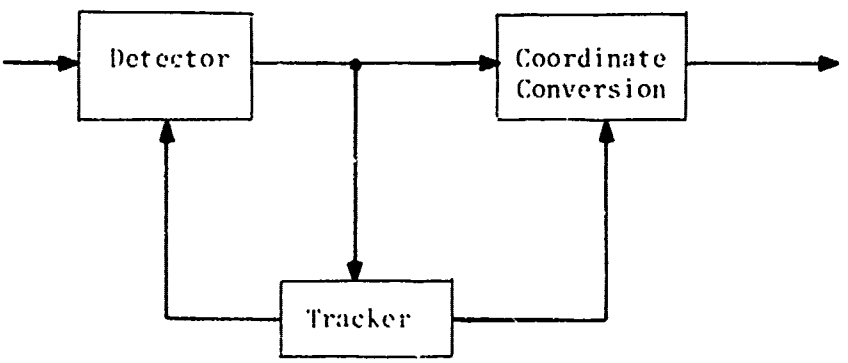
The discussion will assume the use of a predictive filter for tracking. The description of the algorithm and a justification for its use are presented in the first section. The following section discusses the modifications required to make the tracker apply the boundary crossing criteria, and the final section discusses a second technique for collapsing multiple modes by Kalman filtering.

5.1 Considerations of the Tracking Filter

The problem of automatic detection and tracking of targets by the use of HF OTH radar is difficult and complex. Procedures are needed for making detections, converting the coordinates of these detections (usually time delay, doppler shift, and electrical phase difference) to geographical coordinates (range, range rate, and azimuth), and smoothing the resulting tracks. Figure 5-1a shows a block diagram of a straightforward method of carrying out these tasks.

In practice, the system described by Figure 5-1b has the potential to outperform that of Figure 5-1a. The tracker can feed back information to the detector, which will aid the detector by telling it where to look for detections during the next time frame of data. The coordinate conversion procedure can now utilize both the raw and smoothed data in making its decisions and calculations. The tracker will probably require a preliminary type of coordinate conversion on its input data. An additional feedback path from the coordinate conversion block to the tracking block may also be required, although this path is not shown in Figure 5-1b.

(a) System Using No Feedback



(b) System Using Feedback

Figure 5-1. Two Possible System Configurations

The ultimate task is to choose the type of tracking filter which should be utilized and, unavoidably but to a lesser extent, how this filter should interact with the detection and coordinate conversion routines. Choosing a particular type of tracking filter from the wide variety of available filters is not an easy task. Both aerospace and automatic control literature abound with articles on Wiener and Kalman filters (over 200 articles are listed in the bibliography of Kailath, 1974). The alpha-beta filter (Singer and Behnke, 1971) is probably the next most widely discussed filter and, because of its simplicity, has probably been implemented more often in tracking systems than any other filter. A recent report by Cantrell (1975) discusses adaptive versions of alpha-beta and alpha-beta-gamma trackers. Filters which achieve smoothing by fitting polynomials to the observed data have also been proposed (Suyemoto, 1973). Recently, nonlinear adaptive trackers (Bershad, Merryman and Sklansky, 1974) have been investigated; but these results are in the formative stage and have not been considered here.

Singer and Behnke provide a detailed comparison of the Wiener filter, the alpha-beta filter, and the Kalman filter. The relative performance, speed and memory requirements of particular implementations of each of these trackers is discussed. The alpha-beta filter assumes that the target moves in a straight line. Although maneuver detection devices can be used to improve this filter's capability to track maneuvering targets, these devices are probably not applicable to real-time HF radar systems since the detection of maneuvers at a low signal-to-noise ratio and a sampling interval of one second or more, is itself a very difficult problem. Thus, one criterion for an HF radar tracking algorithm should be the ability to follow maneuvering targets without the need for a separate maneuver-detection algorithm.

Two filters which should be considered for the application are the Kalman filter and the Wiener filter. The Wiener filter is actually a simplified type of Kalman filter in which the gain vector (which determines how fast the tracked coordinates can change) does not vary with time and is, in fact, equal to the steady state gain vector of the corresponding Kalman filter. In a Kalman filter, the calculation of the gain vector requires a considerable portion of the total computation time. In the Wiener filter, the gain vector is calculated off-line and is used repeatedly thereafter.

In cases for which the Kalman filter rapidly reaches the steady state, the performance of the Wiener filter will closely approach that of the Kalman filter while requiring about one-fifth the computation time (Singer and Behnke, 1971).

The Kalman filter has an additional advantage over the Wiener filter. In the course of calculating the variable gain vector, the Kalman filter computes an estimate of the covariance matrix of the filter's prediction error. This matrix provides useful information about the probability distribution of the prediction error. If the prediction errors can be assumed to be zero-mean random variables with a jointly Gaussian distribution, the covariance matrix completely specifies their joint distribution function. This distribution function can be fed back to the detector (along with the predicted values) and used to make improved detections on the next frame of data (Oetting, 1974).

For our present purpose, we will assume the more general Kalman filter will be used in the HF radar tracking scheme shown in Figure 5-1b. However, this is done with the reservation that should future testing indicate that the simpler Wiener filter can produce results comparable to those of the Kalman filter, it should be used in order to effect significant savings in computer time and program complexity. A Kalman filter tracking system has been tested and evaluated by Oetting. Insofar as possible, such testing should be made as specific to the situation as is possible to insure a valid comparison of filter performance.

The following material in this section is taken from Oetting (1974) to lay the groundwork for later discussions.

In order to understand the origin of the Kalman filtering equations, we consider a one-dimensional example of target motion, in which only $x_r(t)$, the target's range, is of interest. We can write a Taylor series expansion for $x_r(t)$ in terms of its derivatives as:

$$x_r(t+T) = x_r(t) + T\dot{x}_r(t) + \tfrac{1}{2}T^2\ddot{x}_r(t) + u(t) \qquad (1)$$

where T is an interval of time which is small compared to the reciprocal of the bandwidth of $x_r(t)$, and u(t) is the error term, which results from deviations in the target's acceleration ($\dddot{x}_r \neq 0$). Thus, u(t) is usually referred to as the maneuver noise. The Kalman filter accepts imperfect measurements of $x_r(t)$ and $\dot{x}_r(t)$ and produces smoothed estimates of $x_r(t)$ $\dot{x}_r(t)$, and $\ddot{x}_r(t)$. These measurements are presumed to be corrupted by additive noise, as follows:

$$y_r(t) = x_r(t) + v_r(t) \qquad (2)$$

$$y_{\dot{r}}(t) = \dot{x}_r(t) + v_{\dot{r}}(t)$$

where y_r and $y_{\dot{r}}$ are the measured range and range rate, and v_r and $v_{\dot{r}}$ are the measurement errors.

In order to apply Kalman filter theory to the problem at hand, it is necessary to write the discrete versions of equations (1) and (2):

$$x_r(k+1) = x_r(k) + T\dot{x}_r(k) + \tfrac{1}{2}T^2\ddot{x}_r(k) + u(k) \qquad (3)$$

$$y_r(k) = x_r(k) + v_r(k)$$

$$y_{\dot{r}}(k) = \dot{x}_r(k) + v_{\dot{r}}(k) \qquad (4)$$

where $k = 0, 1, 2, \ldots$ is the index for the sampled values of each time function, and T is now the sampling interval. For the case where u(k), $v_r(k)$, and $v_{\dot{r}}(k)$ are sampled Gaussian white noise processes, it can be

shown that, given the measured quantities y_r and $y_{\dot{r}}$, the Kalman filter calculates estimates for x_r, $x_{\dot{r}}$, and $x_{\ddot{r}}$ (denoted by $\hat{x}_r$, $\hat{x}_{\dot{r}}$, and $\hat{x}_{\ddot{r}}$) which are optimal in the sense of least mean square error, where the range error $e_r = y_r - x_r$ with similar expressions for range rate and acceleration errors.

While the assumption that the measurement noises v_r and $v_{\dot{r}}$ are white and Gaussian is a good assumption in most cases of interest to us, the maneuver noise u will, in general, be neither white nor Gaussian (Kailath, 1974). This situation can be handled in two ways -- we can accept the suboptimality of the Kalman filter for non-white maneuvers, or we can whiten the maneuver noise so that optimal filtering can be achieved. Because of the considerable added computational complexity incurred by the whitening procedure, it may be desirable to accept the suboptimality of the Kalman filter rather than implementing the so-called extended Kalman filter. Of course, in the case where the target exhibits constant radial velocity or constant radial acceleration, the simple Kalman filter will achieve optimality.

5.2 Simplified Two-Dimensional Kalman Filter

We will now discuss the two-dimensional (range and azimuth) Kalman filter which we propose for use in an HF radar system capable of measuring range, doppler, and azimuth. It will be convenient to use matrix notation. Since the independently measureable quantities are range, range rate, and azimuth, the state vector x(k) is given by:

$$x(k) = [x_r(k),\ x_{\dot{r}}(k),\ x_{\ddot{r}}(k),\ x_\theta(k),\ x_{\dot{\theta}}(k)]^t \tag{5}$$

where x_θ and $x_{\dot{\theta}}$ are the azimuth and its time derivative, and the t indicates transpose. The equations of state are:

$$x(k+1) = \Phi x(k) + Gu(k) \tag{6}$$

$$y(k) = Hx(k) + v(k)$$

where $u(k) = [u_r(k),\ u_\theta(k)]^t$ (7)

is the maneuver noise vector,

$$y(k) = [y_r(k),\ y_{\dot{r}}(k), y_\theta(k)]^t \tag{8}$$

is the measurement vector,

$$v(k) = [v_r(k)\ ,\ v_{\dot{r}}(k)\ ,\ v_\theta(k)]^t \tag{9}$$

is the measurement noise vector, and

$$\Phi = \begin{bmatrix} 1 & T & \frac{1}{2}T^2 & 0 & 0 \\ 0 & 1 & T & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & T \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{10}$$

is the transition matrix. The matrices G and H are given by:

$$G = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^t$$

and

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{11}$$

At the K-th sampling instant, the Kalman filter accepts the vector y(k) and recursively calculates four quantities: an estimate of the covariance matrix of the prediction error at time k based on all samples processed prior to time k (denoted by $P(k|k-1)$; a filtered estimate of x(k) based on all samples processed thusfar (denoted by $\hat{x}(k|k)$); an estimate of the covariance matrix of the estimation error based on all samples which have been processed (denoted by $P(k|k)$); and a prediction for x(k+1) based on samples processed through time k (denoted $\hat{x}(k+1)$).

The expressions which must be evaluated (Kailath, 1974) are:

$$P(k|k-1) = \phi P(k-1|k-1)\phi^t + GQG^t \tag{12}$$

$$\hat{x}(k|k) = \hat{x}(k|k-1) + P(k|k-1)H^t W^{-1}(k)\left[y(k) - H\hat{x}(k|k-1)\right] \tag{13}$$

$$P(k|k) = P(k|k-1) - P(k|k-1)H^t W^{-1}(k)HP(k|k-1) \tag{14}$$

$$\hat{x}(k+1|k) = \hat{x}(k|k) \tag{15}$$

In equations (12) - (15), the previously undefined quantities are:

$$Q = \begin{bmatrix} \psi_r^2 & 0 \\ 0 & \psi_\theta^2 \end{bmatrix} \tag{16}$$

$$W(k) - H\,P(k|k-1)H^t + R \tag{17}$$

and

$$R = \begin{bmatrix} \sigma_r^2 & 0 & 0 \\ 0 & \sigma_{\dot{r}}^2 & 0 \\ 0 & 0 & \sigma_\theta^2 \end{bmatrix} \tag{18}$$

where ψ_r^2 and ψ_θ^2 are the variances of the (presumed Gaussian and white) maneuver noises; and σ_r, $\sigma_{\dot{r}}^2$, and σ_θ^2 are the variances of the Gaussian white measurement noises. In order that the Kalman filter be optimal, these variances must be known beforehand; in practice, reasonable values are chosen by experimenting with different combinations of values until good results are obtained in a wide variety of experimental scenarios. Although a dependence on k can be introduced into these parameters, they are indicated as constants in the above equations, since we lack prior knowledge of their possible time variations.

5.3 Modification of the Tracker Algorithm for Mode Identification

From the preceding discussion, we can identify the following information that can be available about each track maintained by the automatic tracker.

a. A set of unsmoothed positions in the radar observables space at which echoes used to develop the track were detected. Each observation comprises the following information: radar delay, doppler shift, azimuth, fine azimuth, signal level (correlator output), frequency of observation and time of the observation.

b. A set of smoothed estimates of position, each estimate comprising the same quantities as those given above.

c. A set of smoothed predictions of future positions in the mode observable region at which echoes are anticipated, each prediction comprising the same quantities as those given above.

d. Estimates of the covariance matrix of the prediction errors and of that of the estimation errors.

Normally, an operating tracker would not maintain all of this information. Instead, it would drop the raw observations $y(k)$ (see Figure 5-2) and retain for display only the set of smoothed estimates of the targets position $\hat{x}(k|k)$ (all k up to present), the most recent prediction of position $\hat{x}(k+1|k)$, and the most recent estimates of the covariance matrix of the estimation error $P(k|k)$.

We now consider the application of the boundary crossing criteria and assume an m of n decision logic is to be used. At first we will consider the information and algorithm required when the mode identifying tracker is to be used with a frequency hopping radar.

Although this mode identifying technique is conceptually quite simple, its implementation becomes somewhat complicated by the need to defer decisions until a sufficient number of observations has been accumulated. The following four entities are required within the data set maintained by the tracker.

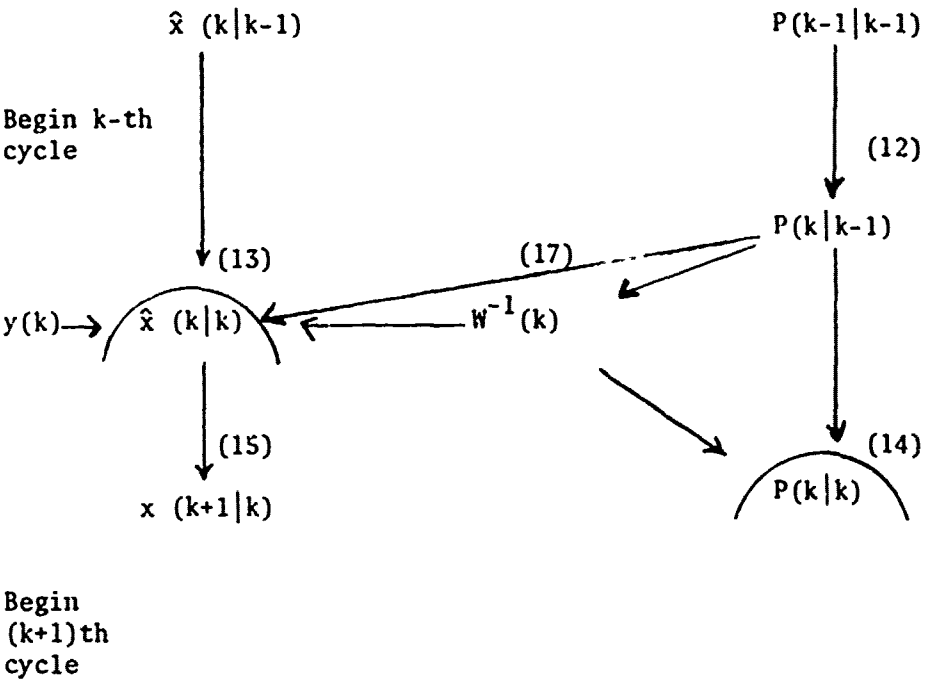


Figure 5-2. Flow of Computation for Kalman Filtering. Numbers in parentheses are equation numbers in text. All other quantities than those shown here are treated as constants.

Item 1

The time delay limits of the various mode observable regions (MOR's) must be recomputed for each frequency. The limits divide the observed range of radar delay into a number of segments. With regard to a particular propagation mode, each segment must be classified as one of three kinds: i.e., as a segment preceding the MOR, within the MOR, and following the MOR. If the MOR is softly bounded, there will also be an interval of uncertainty which for the purpose of the mode identification logic is regarded as yielding no mode identifying information. Note that each segment must be classified as preceding, within, or following the MOR of each possible propagation mode, and that there will in general be several segments falling in each category for any given mode. See Figure 5-3. One set of this information, updated for changed frequency on a dwell-by-dwell basis, is required.

Item 2

For each segment of the radar delay, there must be an identification of its relationship to each MOR. Two bits are required to indicate the relationship of a single segment to a single MOR. This information, like Item 1, can be applied to all tracks and so need only be computed once for each dwell. Furthermore, unless the frequency of the radar changes so much that a boundary for one MOR crosses that of another MOR within the interval of frequency change, the same information can be used from one frequency dwell to the next. For the present, we will assume that the further complication that arises when MOR boundaries cross can be avoided and that the table relating the segments of radar delay to the various MOR's can be computed once for all frequencies of operation.

Item 3

For each track, and for each segment, there must be maintained a record of the successes and failures of attempts to detect the echoes which a track comprises. For the 2 of 5 decision scheme being discussed here, these records must include one bit to indicate prior detections and two bits to keep count of prior "failures-to-detect." See Figure 5-3.

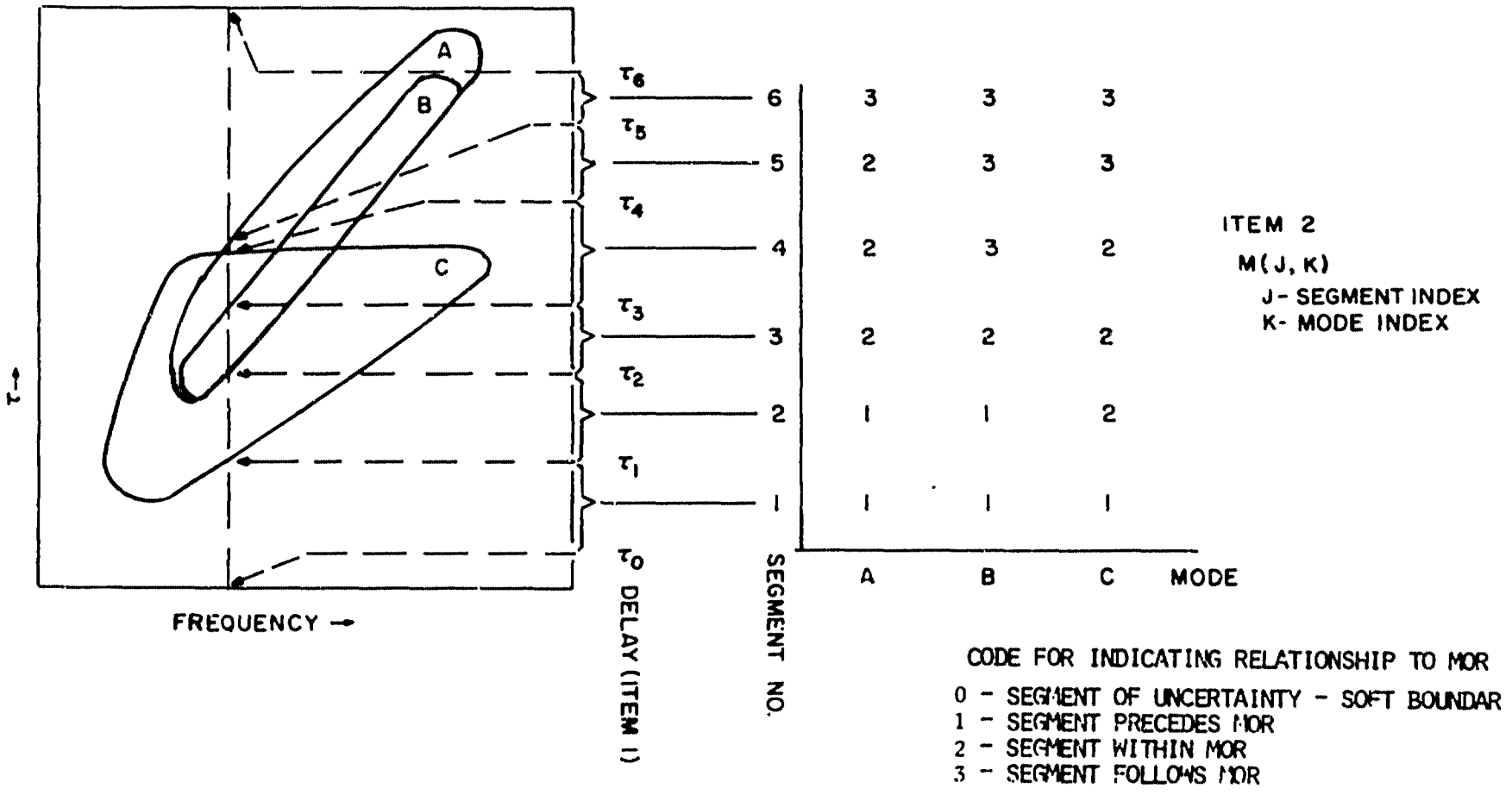


Figure 5-3. Diagram Illustrating Significance of Information Items Carried by Mode Identifying Tracking

Item 4

A summary record of decisions is made on the basis of the above information. Such a record will be carried for each of the segments for each track. Each record will comprise two bits: one bit indicates a decision that echoes have been detected, and the other, that echoes are absent within the associated segment. Thus, we have

0	0	indicate insufficient observations for decision
1	0	M detections have occurred
0	1	n attempts resulted in fewer than n detections
1	1	not used.

to give an indication of mode identity when a decision has successfully been made and of ambiguity prior to identification.

Note that different Items 2, 3, and 4 are required for each track. Item 1 is recomputed on a dwell-by-dwell basis and a single set of these limits applies to all tracks. Furthermore, Items 2 and 3 needto be maintained for a track only until such time that a mode identification has been made, after which Item 4 suffices.

Figure 5-3 shows the required information for mode identification using a 2-of-5 detection scheme. This will be used as an example for the subsequent discussion of the decision algorithm. The need for the various entities will become apparent in the course of the discussion.

The Kalman filter algorithm yields, for each dwell, a predicted position for each of the tracks being maintained. This position would be passed to the automatic detection processor which would search the incoming radar data for an indication of the extension of each track. Because the present considerations of the boundary crossing criteria for mode identification do not involve doppler information the discussion will be limited to the radar delay component of the problem. The automatic tracker searches a neighborhood of the predicted position. For purpose of example, we assume that this neighborhood includes three

range cells beyond, and three before, the predicted position, as well as the predicted cell itself. If a signal exceeding the detection threshold is present in any cell, an additional detection is recorded and used to update the track's smoothed estimate, prediction and covariance matrix. Note that different detection thresholds can be applied to different cells, depending upon their position within the neighborhood.

If a detection is made, its position may be compared with the entries of the limit table (Item 1) to determine in which radar delay segment the detection has occurred. This can be done by comparing the radar delay of the detection with the radar delays listed in Item 2 in order of increasing delay. When a delay which exceeds that of the detection is found, the detection can be assigned to the corresponding interval.

In the event that no detection is made in the course of searching the neighborhood around the predicted position of the echo, it is again necessary to assign a segment to the search so that the count of the number of attempted detections can be augmented. For this purpose the predicted position is compared with the limits of the segments. If the search neighborhood is rather large, it may be desirable to augment the court of attempted detections for all segments that have more than some minimal number of cells within the search neighborhood; however, here we will increment the count only for the region containing the predicted position. Table 5-1 specifies an algorithm for performing the above functions.

As the segments in which a new detection or a failed attempt to detect are identified, the information about the region is updated. In Table 5-1, the index J identifies the segment. The modification of the record (Items 3 and 4) mentioned in Step 6 of Table 5-1 requires only an incrementing of appropriate counters and a test to see whether a decision based on the m of n criterion can be reached. For the purpose of this discussion, we will refer to the record for the Jth segment of radar delay that is maintained for detections which are part of the Ith track by the

Table 5-1. Algorithm for Determining Range Segment of a Detection or of a Search without Detection

N = number of possible modes

(1) If a detection has occurred, set

$D = 1$

$R = \tau$ (radar delay of detection)

If a complete search of the neighborhood indicates no detection set

$D = 0$

$R = P$ (predicted radar delay)

(2) Put counter $J = 1$.

(3) Compare R with limit $\tau(J)$. If $R \leq \tau(J)$ jump to (6).

(4) Increment J. If $J > J_{maximum}$, go to next track.

(5) If $J \leq J_{max}$, jump to (3).

(6) Modify record for segment (J) for track under consideration.

(7) Go to next track.

symbol C(I,J). Also, we make use of the indicator D which is set to 1 if a detection has been made, and 0 if the search of the neighborhood around the predicted position of the detection indicates no detection. Figure 5-3a shows a flow diagram of a subroutine for performing the record updating function.

If updating the record does not lead to a conclusive result, a decision is deferred until the results of more observations are available. If, on the other hand, the new information permits a decision to be made, that the m detections in n attempts either has been or cannot be achieved, a new attempt can be made to identify the propagation mode.

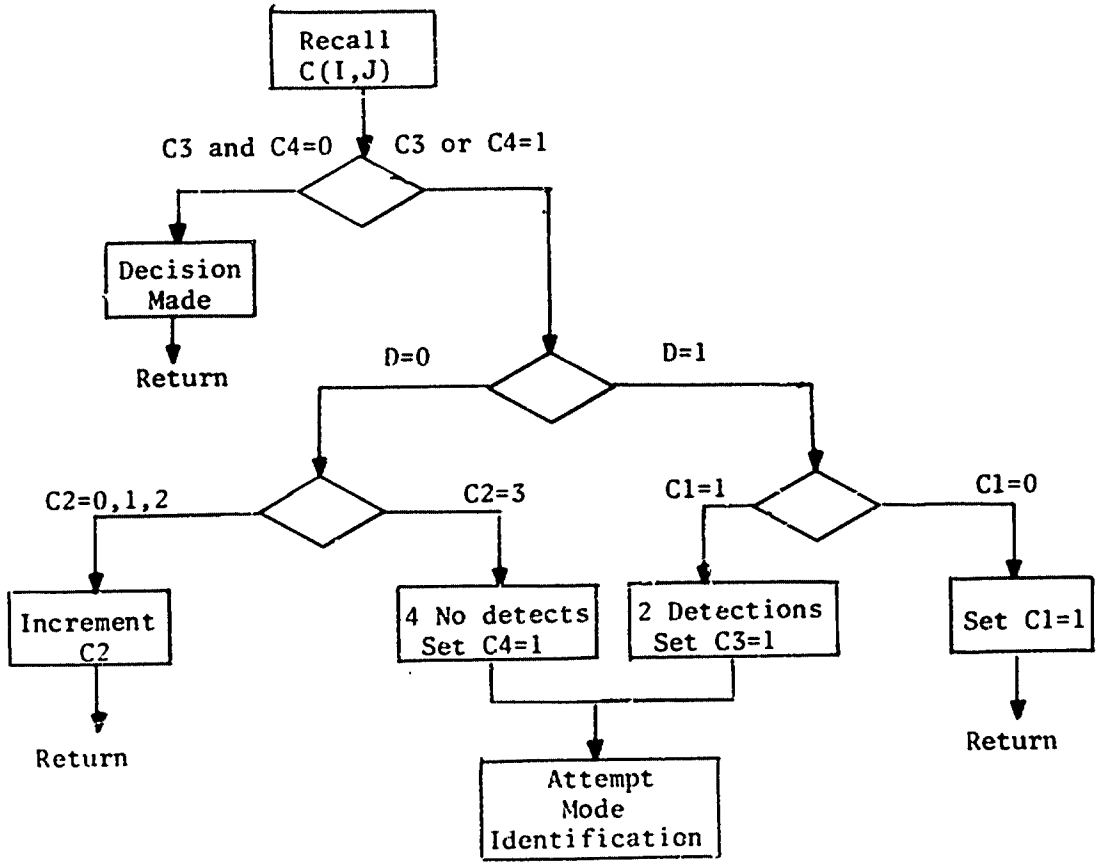
Identification of the propagation mode proceeds in the following stages. First, wherever possible, modes are ruled out, either because a detection (m of n) has occurred in a segment of the radar delay range that is not within the MOR of the eliminated mode, or else because of a failure to detect (m of n) the track within a segment within the MOR. After as many modes as possible are ruled out in this manner, an attempt is made to determine a sole surviving possibility. Note that the segmentation of the radar delay range causes the criterion of termination of a trace at a mode boundary to be applied simultaneously with that involving occurrence of the mode in a forbidden region.

Figure 5-4 is a flow diagram indicating how the data regarding segmental detections and non-detections can be abstracted for use in the decision process. For each track, the records for each segment are examined to determine whether a decision has been made. If so, the segment map (Figure 5-3, Item 2) M(J,K) is examined to determine the implications of a decision. The inference can be that the Kth mode is ruled out or that the detection (or failure to detect) is consistent with the hypothesis that the Kth mode supported the propagation of the echo. The decision that any given mode (e.g., the Kth) can be ruled out is recorded by setting S(K)=1. When the implications of all segmental detections have been drawn regarding all propagation modes, the algorithm goes to the final stage. This consists merely of going through the components of S(K) to determine whether any of them are zero. If only one is zero, the propagation mode is uniquely identified.

Required input arguments D,I,J

C(I,J)

C1	Number of prior detections
C2	Number of prior missed detections
C3	M of N tests complete → detection
C4	M of N tests complete → no detecttion



Algorithm

1. Test C3 and C4. If either is non zero a decision has already been made. Return.
2. If D=1, jump to 5.
3. If C2=3, set C4=1 and jump to 7.
4. Increment C2 and return.
5. If C1=0, set C1=1 and return.
6. Second detection meets 2 of 5 criteria. Set C3.
7. Attempt mode identification with new information (C3 or C4 now equals 1).
8. Return.

Figure 5-3 a. Subroutine for Updating Record C(I,J)

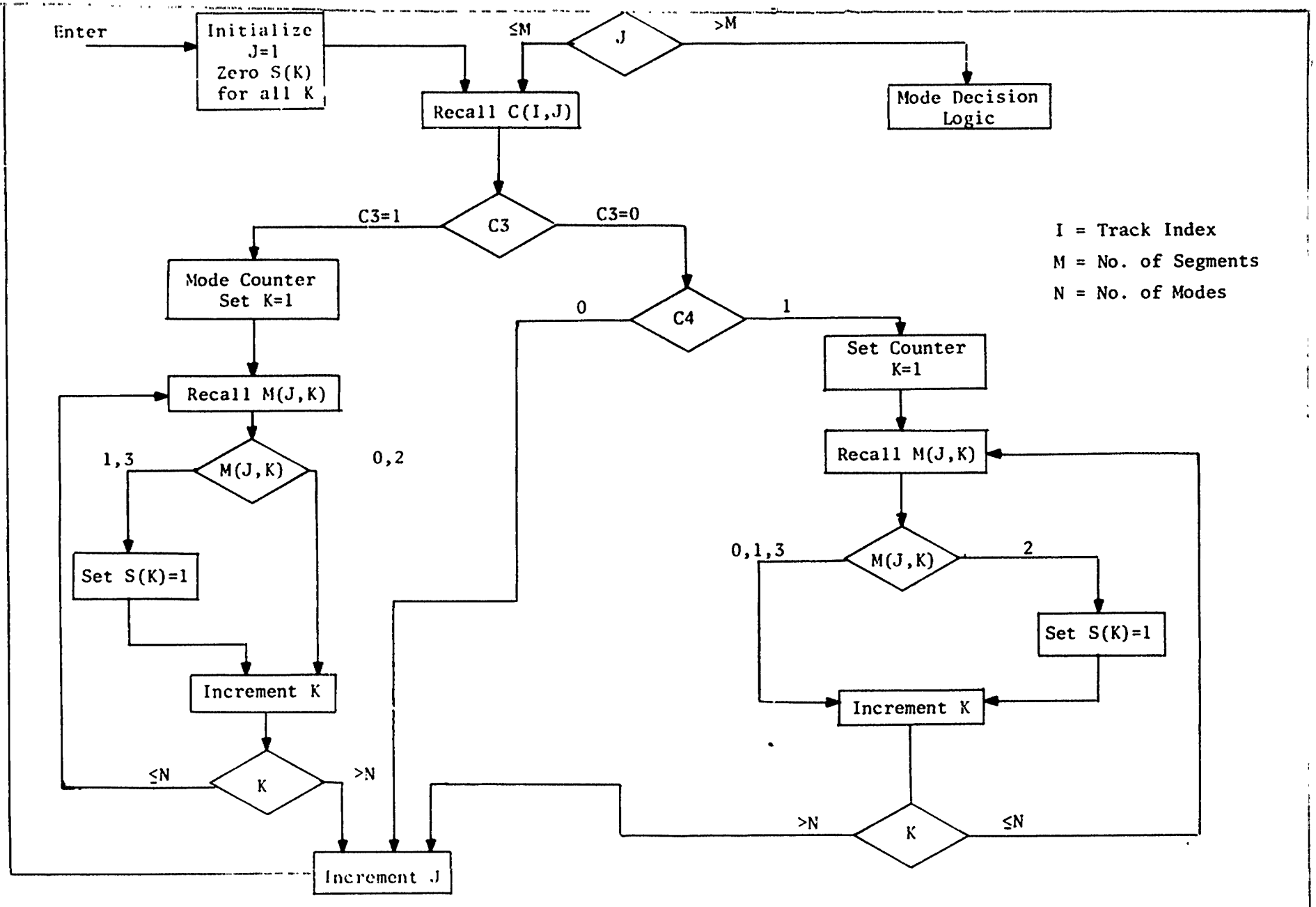


Figure 5-4. Flow Diagram of Abstracting Process for Mode Identification

If more than one is zero, the identity of the mode is still ambiguous, and further attempts at mode identification are necessary. If none is zero, than all modes have been ruled out. This indicates that the trace of the track in the f-τ plane is not consistent with the mode structure assumed to exist. Aside from errors in the assumed mode structure, such effects can occur from errors in the detection process (of either kind), from targets not near the earth's surface, or possibly from ECM spoof targets.

A flow diagram for the final stage of the decision process is shown in Figure 5-5.

5.4 Some Further Considerations

The preceding section discussed the algorithms required to implement the mode identification criteria based on the mode observable regions, as used for a frequency hopping radar. However, one problem was neglected--that of indicating the case in which the frequency changed so much that the boundaries of two MOR's cross within the band of frequency-hopping. Because the identity of segments cannot be preserved across such an intersection of boundaries, a means is required to handle such cases.

Figure 5-6 illustrates an f-τ plane in which such a crossing of boundaries has occurred. The crossing occurs at radar delay τ_c and frequency f_c where the lagging boundary of the B MOR intersects the leading edge of the A MOR. Also included are two mapping arrays, labelled M(I,J) and M'(I,J). The former of these describes the relationship to the MOR's of the various segments defined at frequencies above f_c, where the boundaries cross the latter, at frequencies below f_c. Note that these two mapping arrays differ from one another.

Because the boundaries of MOR's generally slope upward to greater radar delay as frequency is increased, the relationship of the segments into which the radar delay axis is divided at frequencies greater than f_c bear the same relationship to the MOR's as do the segments into which the trace of a target track in the f-τ plane at ranges greater than τ_c -- the radar delay of the boundary intersection. Consequently, for tracks

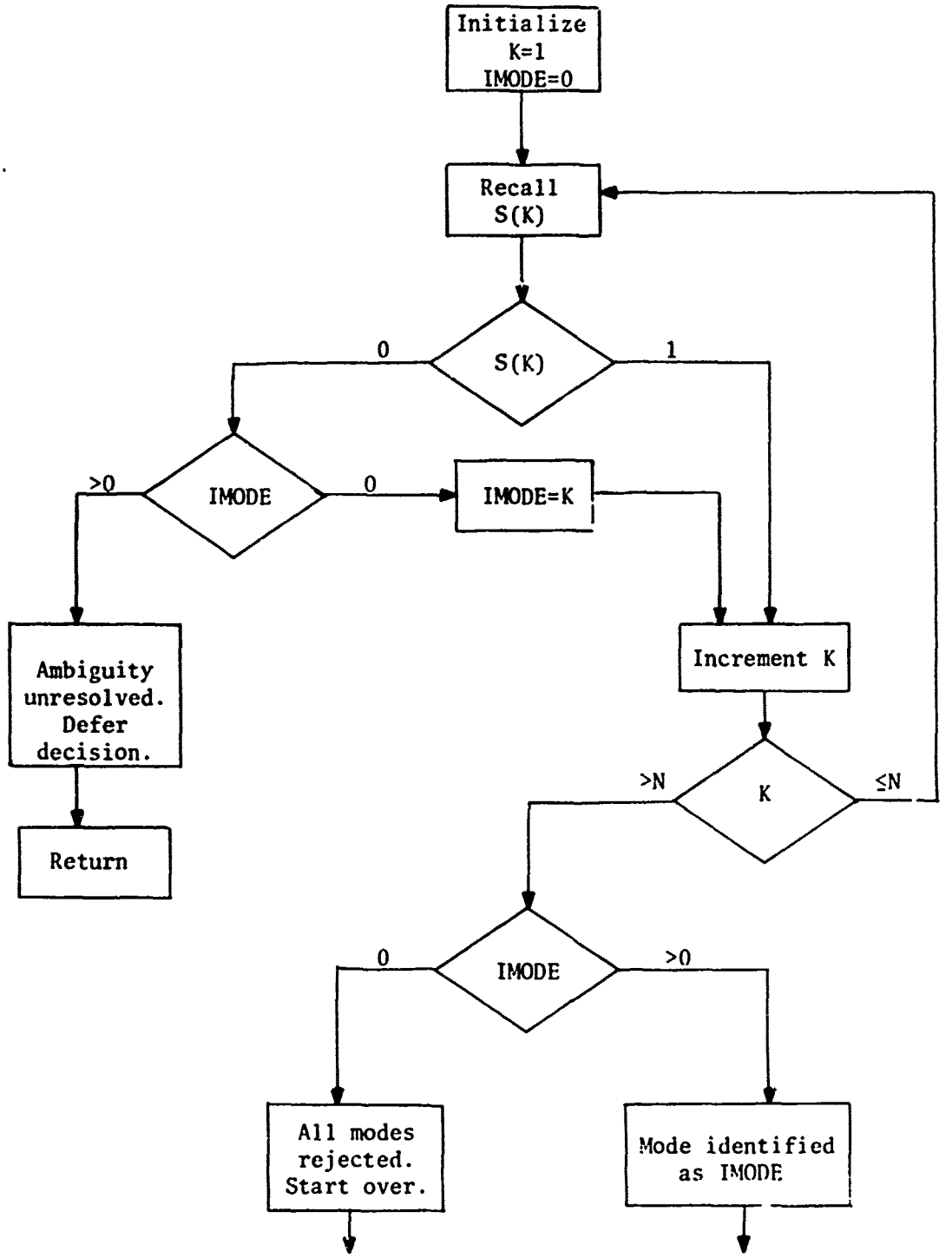
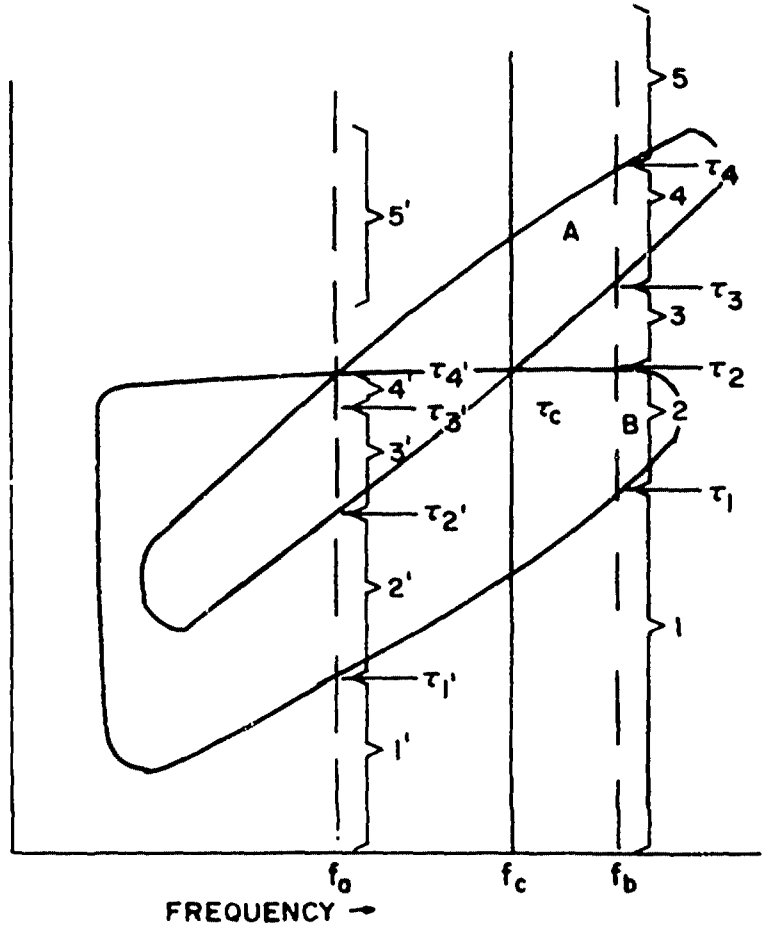


Figure 5-5. Final Mode Identification Process. Three outcomes are possible: (1) ambiguity remains; (2) mode uniquely identified; (3) no mode survived testing. In the last case, tests can be applied again, or the track can be flagged as anomalous and be reported.



M'(I, J)

5	3	3
4	2	3
3	2	2
2	1	2
1	1	1
	A	B

M(I, J)

5	3	3
4	2	3
3	1	3
2	1	2
1	1	1
	A	B

Figure 5-6. Diagram of the f-τ Plane for a Case in which There is a Crossing of MOR Boundaries. Different mapping arrays M'(I,J) and M(I,J) have been defined for use at frequencies below and above the crossing frequency, f_c, respectively.

having a radar delay greater than τ_c, the mapping array M(I,J) should be used, and for tracks with ranges less than τ_c, the mapping array M'(I,J) should be used.

Summarizing, the problem of boundary crossings can be readily handled by defining different mapping arrays to be used with tracks that lie in different intervals of radar delay. In the example shown in Figure 5-6, the array M(I,J) is used for tracks with radar delays exceeding τ_c and M'(I,J), for tracks with radar delays less than τ_c.

5.5 Mode Identification for a Radar Operating at Fixed Frequency

The preceding algorithm assumes that the radar hops from frequency to frequency so that there occur numerous occasions when the extension of the trace into various critical segments can be tested. If these methods are to be applied to a fixed frequency radar, it is necessary to wait for the motion of the radar target to bring it to a position at which the presence or absence of echoes critically indicates the operative mode of propagation. For any given MOR, this happens only twice--at the time the target enters the region the mode covers and when it exits.

Thus, for fixed frequency operation, the mode identification algorithm takes a very simple form: Regions immediately beyond and immediately ahead of each mode's MOR must be continuously monitored for the presence of tracks. Likewise, a similar surveillance must be placed on the portions of the MOR just within each of the two boundaries. If a target is detected outside of the MOR, the corresponding mode is dismissed as a possibility. If it is not detected outside the MOR but is detected within the MOR, then the propagation mode must be that corresponding to the MOR being considered.

In many cases several MOR's have a common boundary, i.e., when modes involved are various combinations of high and low ray modes. In these cases, there is only one critical boundary which, regrettably, is a soft boundary. Consequently, the ability to identify modes in these cases hinge entirely on the radar's ability to definitively determine the presence or absence of echoes in a single critical region.

5.6 Modifications for Soft Boundaries

If a soft boundary occurs, a region must be delineated by a boundary outside of which the presence of echoes can be detected with the required certainty and another outside of which the absence of echoes can be established with the required certainty. The statistics of these determinations is discussed elsewhere, as is the choice of the contours of constant SNR at which these boundaries must be placed.

The newly defined region is one in which the presence or absence of echoes is not a definitive indication of mode identity. These regions can be used to segment the radar delay axis in the same manner as the MOR's were so used. The presence of such segments will be flagged by a zero in the appropriate part of the mapping array M(I,J). The presence of the zero indicates that detections in corresponding segment convey no information regarding the corresponding mode. Figure 5-7 indicates, respectively, the relationship between the distribution of MOR's in the f-τ plane in such cases and the resulting mapping array.

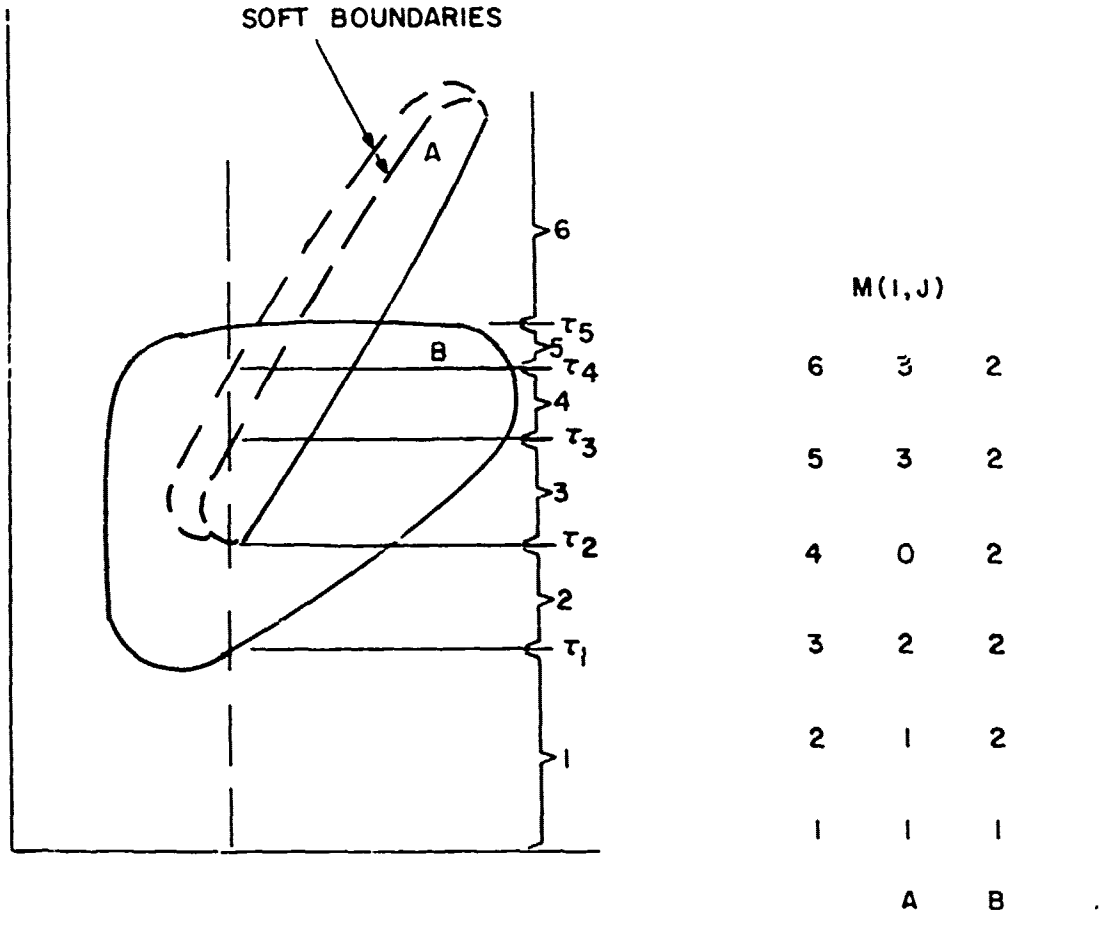


Figure 5-7. The Soft Boundary at the Lagging Edge of the A MOR Leads to the Placement of a 0 in the Segment 4 Position of the Mapping Array M (I,J).

5.7 The Mode Collapsing Process

When it is possible to keep track of the relative distance separating echoes that come from the same target, but via different modes, the possibility exists for the tracker to combine the various echoes. Such a procedure will be discussed in this section. Because the method seems to be particularly attractive for working with high-ray/low-ray combinations of modes, we will concentrate on that case.

At the skip distance the difference in radar delays between the high, low, and combination mode radar echoes is zero. As distance increases beyond the skip distance, the relative delay between the modes increases non-linearly, but monotonically. Let us suppose that the separation in radar delay between the low-ray-mode propagated echo and the mixed-mode-propagated echo is $\Delta(\tau)$ where τ is the radar delay of the low-ray-mode echo. Because of the symmetries of the problem, the high-ray mode will lag the mixed-mode also by $\Delta(\tau)$ and the low-mode by $2\Delta(\tau)$. We also suppose that the monotonic function $\Delta(\tau)$ can be expanded as a power series

$$\Delta(\tau) = \alpha_0 + \alpha_1 \tau + \alpha_2 \tau^2 + \ldots .$$

The α_i's of the above expression are functions of time because the relative delay $\Delta(\tau)$ will vary in its form as the state of the ionosphere changes.

The tracker will be made tolerant to errors in the prediction of range by designing it to search out a finite area of the radar observables space, as we discussed in earlier section. Consequently, the power series in τ can be truncated, so long as the resulting error in $\Delta(\tau)$ is smaller than the tracker's search area. If a reasonably accurate set of estimates of the α's can be maintained, the power series approximation to $\Delta(\tau)$ can be used to predict the radar delay of an echo propagated by a later mode, given that it was detected at a radar delay τ.

The ability to do this permits the effect of multiple propagation modes to be accounted for. Two flow diagrams of variations on the Kalman

filter tracker with provision for multiple modes are shown in Figures 5-8 and 5-9. In Figure 5-8 the Kalman filter used as a tracker is augmented with a second estimator-predictor loop. In this second loop, the expansion coefficients α are estimated. In this scheme the detections made by the later modes do not contribute to the maintenance of the track. Rather, it is assumed that the target can readily be tracked by the first mode. The predicted target state vectors $\hat{x}(k+1|k)$ are then modified to yield predicted state vectors for the later mode echoes, e.g., $\hat{x}'(k+1|k)$, making use of a correction in radar delay calculated from the estimates of the α's, the coefficients of the power series expansion of the mode separation $\Delta(\tau)$. The echoes are then sought in the neightborhood of $\hat{x}(k+1|k)$ and deviations of the observations $y'(k+1)$ are then fed back to the estimator, where the estimates of the coefficients α are updated. The time constants for the estimates of α must be chosen to be relatively long compared to those used in the estimates of the target state vectors $x(k|k)$, so that target prediction errors do not significantly affect the estimates of the α's.

Figure 5-9 diagrams a second method, in which detections made via the later modes are converted to the range of the first mode where they can be included as data points for estimating the track. However, in order to improve the estimates of the α's the observations are separated according to mode and the α's are corrected to remove any observed bias.

This scheme provides a greater number of data points with which the tracker can work but does not improve the ability to estimate the α's. These are based upon a division of the available data points into three groups, according to propagation mode. However if as before the estimates of the α's are based on a longer history of observations than are those of the tracks, the time averaging can more than make up for the effect of the division of the data points into groups according to propagation mode.

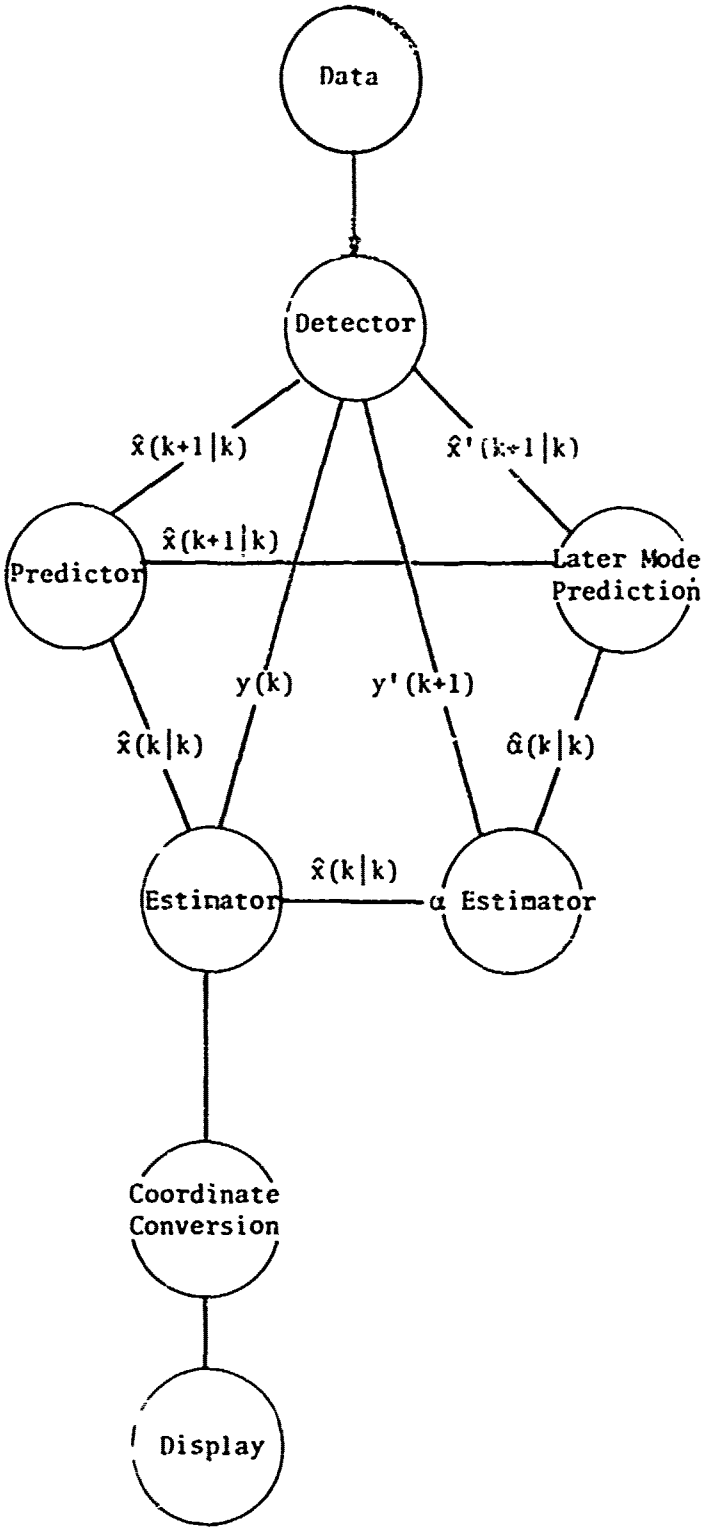


Figure 5-8. Modification of the Tracker Algorithm to Extract Multiple Mode Effects.

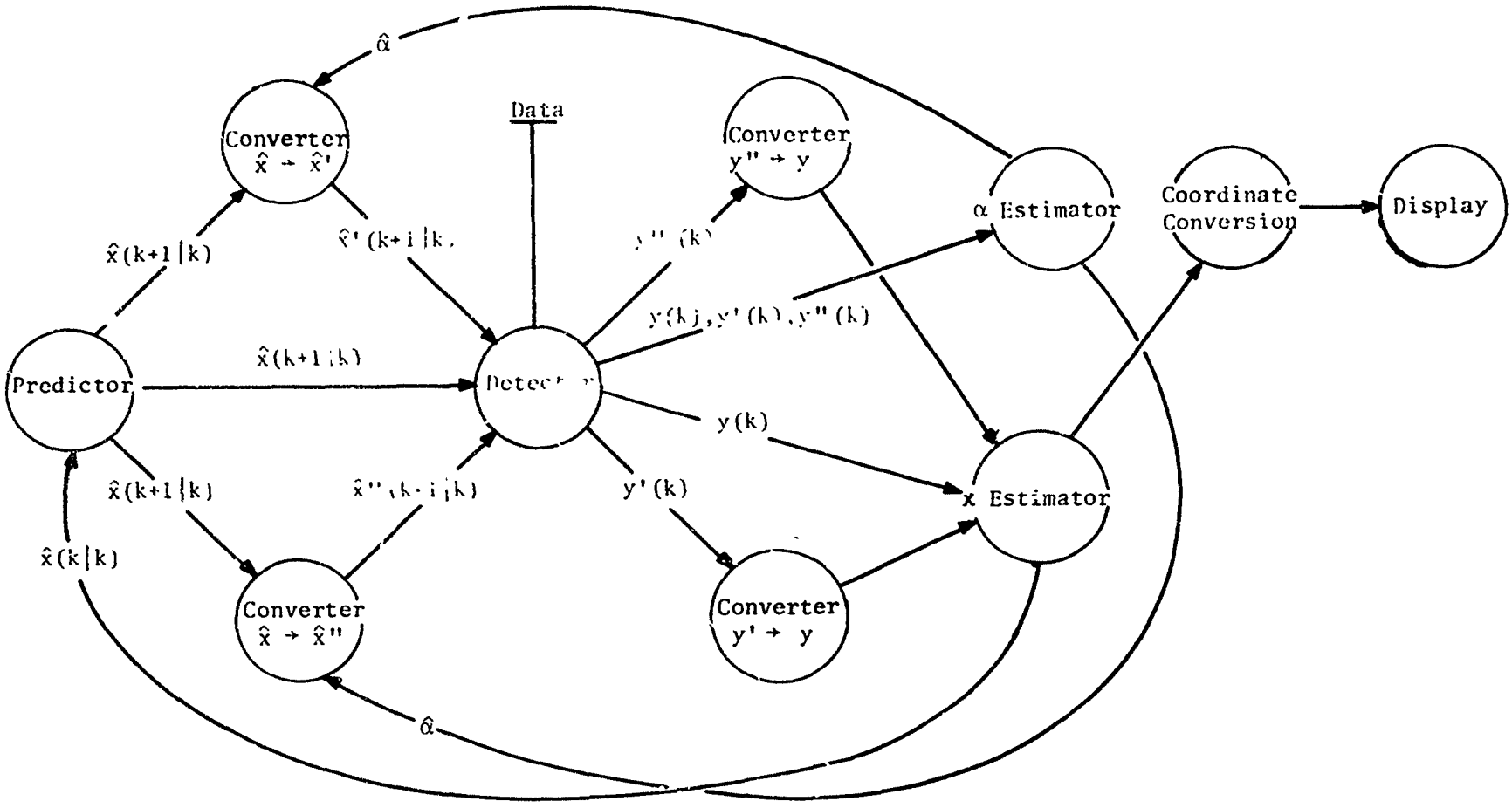


Figure 5-9. Flow Diagram of Tracker Which Tracks Mixed Mode Detections.

A further note regarding the operation of the tracker in Figure 5-9 concerns the relative signal strengths of the echoes propagated via the different modes. Usually, one of the modes will provide less path loss than does the other, with the result that there may be substantial differences in the signal-to-noise ratio of echoes received by one mode and those of another. Consequently, the probability of detection and false alarm rate cannot be kept the same for the detection of echoes propagated by the different modes. This difference would have to be accounted for in the tracker so that, for example, the tracker would not drop track on the basis of having failed to detect the echo as propagated by a particularly weak mode.

The above discussion has concentrated on the radar delay of the observed targets. However, a similar set of conditions apply to the doppler shift as well. Just as was true for the radar delay, there will be a difference, from one mode to another, of the doppler shift observed on the echoes. Again, these can be expanded as a function of radar delay and used to interrelate the state vectors describing the echo as observed by the different modes.

5.8 References

1. Bershad, N.J., P. Merryman and T. Sklansky, "Adaptive Trackers Based on Continuous Learning Theory," IEEE Transactions on Aerospace and Electronic Systems, March 1974, pp. 245-253.

2. Cantrell, B.H., "Adaptive Tracking Algorithm," NRL Memorandum Report 3037, Naval Research Laboratory, Washington, D.C., April 1975.

3. Kailath, T., "A View of Three Decades of Linear Filtering Theory," IEEE Transactions on Information Theory, March 1974, pp. 146-181.

4. Oetting, J., "Application of Kalman Filtering to Automatic Tracking and Smoothing," Internal Technical Memorandum No. 337, ITT Electro-Physics Laboratories, Inc., Columbia, Maryland, September 1974.

5. Singer, R.A., and K.W. Behnke, "Real-Time Tracking Filter Evaluation and Selection for Tactical Applications," IEEE Transactions on Aerospace and Electronic Systems, January 1971, pp. 100-110.

6. Suyemoto, L., "Summary of Results from Contractors' Tracking Studies for an OTHB Radar (U)," report prepared by Mitre Corporation for Deputy for Surveillance and Control Systems, Electronic Systems Division, Air Force Systems Command, Report No. ESD-TR-73-139, April 1973 (SECRET).

6. CONCLUSIONS

The existence of multiple propagation modes can have two kinds of effects on the data developed and displayed by an OTH radar. The first effect is to introduce an ambiguity in the interpretation of tracks. Correct identification of the propagation mode is necessary before the observed radar delay of the echo can be interpreted in terms of a distance from the radar. Misidentification can result in an error amounting to several tens to hundreds of km in the position assigned to an observed track. Because echoes propagated by different modes can occur at a common range, the risk of making such an error exists.

The second effect is to cause a single aircraft, or other target, to produce multiple echoes at the radar, each having propagated by a different propagation mode or combination of such modes, and consequently, each arriving with a different radar delay. If the common origin of these echoes is not recognized, the radarist will overestimate the number of targets he is viewing. Conversely, if he wrongly associates echoes, assuming them to have been caused by multimoding when, indeed, they have not been, he will underestimate the number of targets.

Multimoding between distinct propagation modes can occur during the daytime when the F layer is accompanied by sunlight sensitive underlying layers, namely the E and F1 layers. It can also occur during summer nights when the sporadic E layer occurs. However, confusion of E-and F-layer propagated echoes is only a problem at the near ranges reached by the relatively short distance E layer.

The risk of confusion also exists between one-hop and two-hop modes of a single layer. These can usually be avoided if the operating frequency of the radar can be made to exceed 75% of the maximum observable frequency (MOF) at the range of interest.

Finally, the confusion of high-ray and low-ray propagated echoes, and combinations of these, can also lead to problems. Because of

defocussing effects, the high ray becomes significantly attenuated at frequencies more than a couple MHz below the MOF to the range of interest. Unfortunately, the obscuring effect of underlying layers may preclude such a choice of operating frequency.

A number of tests can be applied to determine the propagation mode responsible for observed track. Many of these involve knowledge of the maximum observable frequency to the range in question and to other details of the propagation conditions, such as the maximum radar delay at which echoes propagated via a particular mode can appear.

To facilitate the discussion of these mode identifying tests, a number of concepts were introduced. Among these were the mode observable region (MOR) for a given propagation mode, being the region in the frequency-delay plane where surface targets (which here includes low aircraft) can occur. The boundaries of the mode observable regions can be discretely bounded or can be the consequence of a gradual diminution of signal strength. The adjectives "hard" and "soft" were applied respectively to such boundaries.

It was shown that the detectability of echoes on either side of such boundaries provided a means for testing the identity of the propagation mode responsible for any given track. The statistical considerations required for such a determination at hard and soft boundaries was related to the maximum acceptable probability of misidentifying the propagation mode.

Various tracker algorithms were considered, and particular attention was given to the suitability of a Kalman filter tracker for incorporation into a mode identifying tracker. To be suitable, it is necessary that the associated automatic detection scheme be capable of producing a null decision.

Algorithms were provided for applying the mode identifying criteria to tracks as developed by a Kalman filter. As an example, the case where the required certainty of mode identification would be achieved with a 2-out-of-5 detection logic was considered.

Additional flow diagrams are presented for modifying the Kalman filter to perform a mode-collapsing process whereby multiple tracks produced by a single target could be associated by maintaining an estimate of mode separation as a function of radar delay.

This latter process has particular value for the case when the multimoding is the result of high-ray, low-ray and combination modes. In such a case, the different modes have one boundary in common, while the other boundaries are soft, making the application of the boundary-crossing criteria difficult.